

UNITED STATES AIR FORCE
SUMMER RESEARCH PROGRAM -- 1996
SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 5C

WRIGHT LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscas

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1996

ARM 01-86-1262

20010321 052

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paper Project Collection (0734)

viewing
ormation

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December, 1996		3. REPORT NUMBER	
4. TITLE AND SUBTITLE 1996 Summer Research Program (SRP), Summer Faculty Research Program (SFRP), Final Reports, Volume 5C, Wright Laboratory				5. FUNDING NUMBERS F49620-93-C-0063	
6. AUTHOR(S) Gary Moore					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research & Development Laboratories (RDL) 5800 Uplander Way Culver City, CA 90230-6608				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) 801 N. Randolph St. Arlington, VA 22203-1977				10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES					
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the SFRP participants at the Wright Laboratory.					
14. SUBJECT TERMS Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student				15. NUMBER OF PAGES	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL		

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to **stay within the lines** to meet **optical scanning requirements**.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available
(e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract
G - Grant
PE - Program
Element

PR - Project
TA - Task
WU - Work Unit
Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es).
Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es).
Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.

Leave blank.

NASA - Leave blank.

NTIS -

Block 13. Abstract. Include a brief (*Maximum 200 words*) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (*NTIS only*).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 5 is bound in three parts, 5A, 5B and 5C. Volume 5A contains #1-24. Volume 5B contains reports #25-48 and 5C contains #49-70. The Table of Contents for Volume 5 is included in all parts.

This document is one of a set of 16 volumes describing the 1996 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME

TITLE

1	Program Management Report
<i>Summer Faculty Research Program (SFRP) Reports</i>	
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A , 5B & 5C	Wright Laboratory
6	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
<i>Graduate Student Research Program (GSRP) Reports</i>	
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center, and Wright Patterson Medical Center
<i>High School Apprenticeship Program (HSAP) Reports</i>	
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

SFRP FINAL REPORT TABLE OF CONTENTS

i-xii

1. INTRODUCTION	1
2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM	2
3. RECRUITING AND SELECTION	3
4. SITE VISITS	4
5. HBCU/MI PARTICIPATION	4
6. SRP FUNDING SOURCES	5
7. COMPENSATION FOR PARTICIPATIONS	5
8. CONTENTS OF THE 1996 REPORT	6

APPENDICIES:

A. PROGRAM STATISTICAL SUMMARY	A-1
B. SRP EVALUATION RESPONSES	B-1

SFRP FINAL REPORTS

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Richelle M Allen-King	Washington State University , Pullman , WA Reduction Kinetics in a Batch Metallic Iron/Water System:Effect of Iron/Water Exposure	AL/EQC	2- 1
DR Anthony R Andrews	Ohio University , Athens , OH Investigation of the Electrochemiluminescent Properties of Several Natural & Synthetic Compounds	AL/EQC	2- 2
DR Deborah L Armstrong	Univ of Texas at San Antonio , San Antonio , TX Development of A primary Cell Culture Preparation for Studying Mechanisms Governi ng Circadian Rhyth	AL/CFTO	2- 3
DR Robert L Armstrong	New Mexico State University , Las Cruces , NM Microparticle Bioluminescence	AL/CFD	2- 4
DR Maureen E Bronson	Wilkes Univ School of Pharmacy , Wilkes-Barre , PA Lack of Effect of UltraWideband Radiation on Pentylenetetrazol-Induced Convulsions in Rats	AL/OER	2- 5
DR Marc L Carter, PhD, PA	University of South Florida , Tampa , FL Assessment of the Reliability of Ground-Based Observers for the Detection of Aircraft	AL/OEO	2- 6
DR Jer-Sen Chen	Wright State University , Dayton , OH A Study of Data Compression Based on Human Visual Perception	AL/CFHV	2- 7
DR Cheng Cheng	Johns Hopkins University , Baltimore , MD Sequential Optimization Algorithm for Personnel Assignmt Based on Cut-Off Profiles & Rev of Brogden	AL/HRM	2- 8
DR Elizabeth T Davis	Georgia Institute of Tech , Atlanta , GA Perceptual Issues in Virtual Environments & Other Simulated Displays	AL/CFHP	2- 9
DR Keith F Eckerman	Univ of Tennessee , Knoxville , TN	AL/OEB	2- 10
DR Paul A Edwards	Edinboro Univ of Pennsylvania , Edinboro , PA A Viartion Fuel Identification- Neural Network Analysis of the Concentration of Benzene and Naphtha	AL/EQC	2- 11

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Randolph D Glickman	Univ of Texas Health Science Center , San Antonio , TX A Study of Oxidative Reactions Mediated by Laser-Excited Ocular Melanin	AL/OEO	2- 12
DR Ellen L Glickman-Weiss	Kent State University , Kent , OH The Effect of Short Duration Respiratory Musculature Training on Tactical Air Combat	AL/CFTF	2- 13
DR Irwin S Goldberg	St. Mary's Univ of San Antonio , San Antonio , TX Development of a Physiologically-Based Pharmacokinetic Model for the Uptake of Volatile Chemicals during	AL/OES	2- 14
DR Robert J Hirko	University of Florida , Gainesville , FL Investigation of The Suitability of Tactile and Auditory Stimuli for use in Brain Actuated Control	AL/CFHP	2- 15
ISU VPP Acct4212313(Dooley)	Iowa State University , Ames , IA Determination of the Influence of Ultrawideband Exposure of Rats During Early Pregnancy on Pregnancy	AL/OER	2- 16
DR Andrew E Jackson	Arizona State University , Tempe , AZ A Description of Integrated Joint Use Initiatives to Satisfy Customer Requirements Across Govt Academia	AL/HRA	2- 17
DR John E Kalns	Ohio State University , Columbus , OH	AL/AOHR	2- 18
DR Nandini Kannan	Univ of Texas at San Antonio , San Antonio , TX Modeling Decompression Sickness Using Survival Analysis Techniques	AL/CFTS	2- 19
DR Antti J Koivo	Purdue Research Foundation , West Lafayette , IN Skill Evaluation of Human Operators	AL/CFBA	2- 20
DR Suk B Kong	Incarnate Word College , San Antonio , TX Aromatic Hydrocarbon Components in Diesel, Jet-A And JP-8 Fuels	AL/OEA	2- 21
DR Xuan Kong	Northern Illinois University , De Kalb , IL Mental Workload Classification via Physiological Signal Processing: EOG & EEG Analyses	AL/CFHP	2- 22

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Charles S Lessard	Texas A & M Univ-College Station , College Station , TX Preliminary Studies of Human Electroencephalogram (EEG) Correlates of GzAcceleration Tolerance	AL/CFTO	2- 23
DR Audrey D Levine	Utah State University , Logan , UT Biogeochemical Assessment of Natural Attenuation of JP-4 Contaminated Ground Water	AL/EQC	2- 24
DR David A Ludwig	Univ of N.C. at Greensboro , Greensboro , NC The Illusion of Control & Precision Associated w/Baseline Comparisons	AL/AOCY	2- 25
DR Robert G Main	Cal State Univ, Chico , Chico , CA Designing Instruction For Distance Learning	AL/HRT	2- 26
DR Phillip H Marshall	Texas Tech University , Lubbock , TX Time to Contact Judgments in The Presence of Static and Dynamic Objects: A Preliminary Report	AL/HRM	2- 27
MS Sandra L McAlister	Stonehill College , North Easton , MA	AL/AO	2- 28
MR Bruce V Mutter	Bluefield State College , Bluefield , WV Environmental Cost Analysis: Calculating Return on Investment for Emerging Technologies	AL/EQP	2- 29
DR Sundaram Narayanan	Wright State University , Dayton , OH Java-Based Application of the Model-View-Controller Framework in Developing Interfaces to interactive	AL/HRT	2- 30
DR Karl A Perusich	Purdue University , South Bend , IN Examining Alternate Entry Points in a Problem Using Fuzzy Cognitive Maps	AL/CFHI	2- 31
DR Judy L Ratliff	Murray State Univ , Murray , KY A Study of The Ability of Tunicates to be used as Global Bioindicators	AL/EQC	2- 32
DR Paul D Retzlaff	Univ of Northern Colorado , Greeley , CO Computerized Neuropsychological Assessment of USAF Pilots	AL/AOCN	2- 33

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR William G Rixey	University of Houston , Houston , TX The use of Solid-Phase Microextraction (SPME) for the low level Detection of BTEX and PAHs In Aqueou	AL/EQC	2- 34
DR Ali M Sadegh	CUNY-City College , New York , NY Investigation of Neck Models for Predicting Human Tolerance to Accelerations	AL/CFBE	2- 35
DR Kandasamy Selvavel	Claflin College , Orangeburg , SC Truncated Bivariate Exponential Models	AL/AOEP	2- 36
DR Barth F Smets	University of Connecticut , Storrs , CT Biodegradation of 2-4-DNTand 2,6-DNT in Mixed Culture Aerobic Fluidized Bed Reactor and Chemostat	AL/EQC	2- 37
DR Mary Alice Smith	University of Georgia , Athens , GA A Study of Apoptosis During Limb Development	AL/OET	2- 38
DR Daniel P Smith	Utah State University , Logan , UT Bioremediation & its Effect on Toxicity	AL/EQW	2- 39
MR. Joseph M Stauffer	Indiana State University , Terre Haute , IN Joint Corrections for Correlation Coefficients	AL/HRMA	2- 40
DR William B Stavinoha	Univ of Texas Health Science Center , San Antonio , TX Studies to Identify Characterisitic Changes in the Urine Following Ingestion of Poppy seed	AL/AOT	2- 41
DR William A Stock	Arizona State University , Tempe , AZ Application of Meta-Analysis to Research on Pilot Training	AL/HRA	2- 42
DR Nancy J Stone	Creighton University , Omaha , NE Engagement, Involvement, and Self-Regualted Leearnign Construct and Measurement Development to Asses	AL/HRT	2- 43
DR Brenda M Sugrue	Univ of Northern Colorado , Greeley , CO Aptitude-Attribute Interactions in Test Performance	AL/HRTI	2- 44

SRP Final Report Table of Contents

Author	University/Institution	Armstrong Laboratory	Vol-Page
Report Title	Directorate		
DR Stephen A Truhon	Winston-Salem State University, Winston-Salem, NC	AL/HRM	2 - 45
	Mechanical Specialties in the U.S. Air Force: Accession Quality & Selection Test Validity		
DR Mariusz Ziejewski	North Dakota State University, Fargo, ND	AL/CFBV	2 - 46
	Validation of the Deformable Neck Model for A +Gz Acceleration		

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Graham R Allan	New Mexico Highlands University, Las Vegas, NM Temporal and Spatial Characterization of a Synchronously-Pumped Periodically-Poled Lithium Niobate Optical	PL/LIDN	3 - 1
DR Brian P Beecken	Bethel College, St. Paul, MN Testing of a Dual-Band Infrared Focal Plane Array & An Infrared Camera Sys	PL/VTRP	3 - 2
DR Mikhail S Belen'kii	Georgia Inst of Technology, Atlanta, GA Tilt Sensing Technique w/Small Aperture Beam & Related Physical Phenomena	PL/LIG	3 - 3
DR Asoke K Bhattacharyya	Lincoln University, Jefferson City, MO Part A: Effect of Earth's Surface & Loss on the Resonant Frequencies of Buried Objects	PL/WSQ	3 - 4
DR Joseph M Calo	Brown University, Providence, RI Transient Studies of the Effects of Fire Suppressants in a Well-Stirred Combustor	PL/GPID	3 - 5
DR James J Carroll	Youngstown State University, Youngstown, OH Examination of Critical Issues in the use of (178) hf For High Energy Density Applications	PL/WSQ	3 - 6
DR Soyoung S Cha	Univ of Illinois at Chicago, Chicago, IL A Study on Hartmann Sensor Application to Flow Aero-Optics Investigation Through Tomographic Recons	PL/LIMS	3 - 7
DR Tsuchin Chu	Southern Illinois Univ-Carbondale, Carbondale, IL	PL/RKS	3 - 8
DR Kenneth Davies	Univ of Colorado at Boulder, Boulder, CO Studies of Ionospheric Electron contents and High-Frequency Radio Propagation	PL/GPIM	3 - 9
DR Judith E Dayhoff	Univ of Maryland, College Park, MD Dynamic Neural Networks: Prediction of an Air Jet Flowfield	PL/LIMS	3 - 10
DR Ronald R DeLyser	University of Denver, Denver, CO Analysis of Complex Cavities Using the Finite Difference Time Domain Method	PL/WSTS	3 - 11
DR Andrew G Detwiler	S Dakota School of Mines/Tech, Rapid City, SD Evaluation of Engine-Related Factors Influencing Contrail Prediction	PL/GPAB	3 - 12
DR Itzhak Dotan	The Open University of Israel, Tel-Aviv Israel Studies of Ion-Molecule Reaction Rates at Very High Temperatures	PL/GPID	3 - 13

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Omar S Es-Said	Loyola Marymount University, Los Angeles, CA On the Matis Selection of Durable Coatings for Cryogenic Engineer Technology	PL/RKE	3 - 14
DR Jeffrey F Friedman	University of Puerto Rico, Mayaguez, PR Testing the Frozen Screen Model of Atmospheric Turbulence	PL/LIMI	3 - 15
DR John A Guthrie	University of Central Oklahoma, Edmond, OK Ultrawide-Band Microwave Effects Testing on an Electronic System	PL/WSMA	3 - 16
DR George W Hanson	Univ of Wisconsin - Milwaukee, WI A Volumetric Eigenmode Expansion Method for Dielectric Bodies	PL/WSQ	3 - 17
DR Mayer Humi	Worcester Polytechnic Inst., Worcester, MA Wavelets and Their Applications to the Analysis of Meteorological Data	PL/GPAA	3 - 18
DR Christopher H Jenkins	S Dakota School of Mines/Tec, Rapid City, SD Shape Control of An Inflated Thin Circular Disk	PL/VT	3 - 19
DR Dikshitulu K Kalluri	University of Lowell, Lowell, MA Electromagnetic Wave Transformation in a Two-Dimensional-Space-Varying and Time-Varying Magnetoplasma	PL/GPIA	3 - 20
DR Aravinda Kar	University of Central Florida, Orlando, FL Thick Section Cutting w/Chemical Oxygen-Iodine Laser & Scaling Laws	PL/LIDB	3 - 21
DR Spencer P Kuo	Polytechnic University, Farmingdale, NY Theory of Electron Acceleration by HF-Excited Langmuir Waves	PL/GPI	3 - 23
DR Andre Y Lee	Michigan State University, East Lansing, MI Characterization Methods for Adhesion Strength Between Polymers & Ceramics	PL/RKS	3 - 24
DR Bruce W Liby	Manhattan College, Riverdale, NY Acousto-Optic Retro-Modulator	PL/VTRA	3 - 25
DR Feng-Bao Lin	Polytechnic Inst of New York, Brooklyn, NY Structural Ballistic Risk Assessment-Fracture Modeling	PL/RKEM	3 - 26
DR M Arfin K Lodhi	Texas Tech University, Lubbock, TX Theory, Modeling & Analysis of AMTEC	PL/VTP	3 - 27

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Ronald A Madler	Embry-Riddle Aeronautical University, Prescott, AZ Estimating the Area of Artificial Space Debris	PL/WSAT	3 - 28
DR Carlos A Ordonez	University of North Texas, Denton, TX Boundary Conditions at A Plasma-Facing Surface	PL/WSQA	3 - 29
DR Michael J Pangia	Georgia Southwestern Coll, Americus, GA Further Analysis of Kilohertz Order Waves Associated with Electron Beam Operations on STS46	PL/GPSG	3 - 30
DR Ronald M Pickett	University of Lowell, Lowell, MA Temporal-Displacement Stereograms of the Ionosphere: An Exploration of Their Utility in the Analysis of Equatorial Emission Depletion Bands	PL/GPIA	3 - 31
DR Edgar Sanchez-Sinencio	Texas A&M Univ-College Station, College Station, TX Low Voltage Analog Circuit Design for Radiation Tolerance	PL/VTER	3 - 32
DR Joseph C Slater	Wright State University, Dayton, OH Smart Structure/Actuator Modeling 7 Design for the Integrated Ground Demonstration Lab	PL/VTI	3 - 33
DR Ashok Srivastava	Louisiana State University, Baton Rouge, LA Modeling of Total Dose Response of SOI N-MOSFETS for Low Power CMOS Circuits	PL/VTER	3 - 34
DR James M Stiles	University of Kansas, Lawrence, KS The Potential Applications of Super-Resolution & Array Processing to Space-Based Radars	PL/VTRA	3 - 35
DR Charles M Swenson	Utah State University, Logan, UT Balloon Launch Retromodulator Experiment	PL/VTRA	3 - 36
DR Miguel Velez-Reyes	University of Puerto Rico, Mayaguez, PR Regularization Methods for Linear and Nonlinear Retrieval Problems	PL/GPAS	3 - 37

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
DR A F Anwar	University of Connecticut, Storrs, CT A Study of Quantum Wells Formed in Al _x Ga _{1-x} As _{1-y} In _z Ga _{1-z} As/Al _x Ga _{1-x} As _{1-y} Heterostructures	RL/ER	4 - 1
DR Ercument Arvas	Syracuse University, Syracuse, NY An Assessment of the Current State of the Art of Stap from an Electromagnetics Point of View	RL/OCSS	4 - 2
DR Ahmed E Barbour	Georgia Southern University, Statesboro, GA Formal Verification Using ORA Larch/VHDL Theorem Prover	RL/ERDD	4 - 3
DR Milica Barjaktarovic	Wilkes University, Wilkes Barre, PA Formal Specification and Verification of Missi Architecture Using Spin	RL/C3AB	4 - 4
DR Daniel C Bukofzer	Cal State Univ, Fresno, Fresno, CA Performance Analysis & Simulation Results of Delay & Spread Spectrum Modulated Flip Wave-Signal Gene	RL/C3BA	4 - 5
DR Xuesheng Chen	Wheaton College, Norton, MA Optical and Non-Destructive Methods to Determine the Composition and Thickness of an In _x Ga _{1-x} As/InP	RL/ERX	4 - 6
DR Jun Chen	Rochester Inst of Technology, Rochester, NY A Study of Optoelectronic Feedback-Sustained Pulsation of Laser Diodes at 1300 nm & 780 nm	RL/OCPA	4 - 7
DR Everett E Crisman	Brown University, Providence, RI Evaluation of Semiconductor Configurations as Sources for Optically Induced Microwave Pulses	RL/ERAC	4 - 8
DR Digendra K Das	SUNYIT, Utica, NY Techniques for Determining of the Precision of Reliability Predictions and Assessments.	RL/ERSR	4 - 9
DR Matthew E Edwards	Spelman College, Atlanta, Ga The Analysis of PROFILER for Modeling the Diffusion of Aluminum-Copper on a Silicon Substrate	RL/ERDR	4 - 10
DR Kaliappan Gopalan	Purdue University - Calumet, Hammond, IN Speaker Identification & Analysis of Stressed Speech	RL/IRAA	4 - 11
DR Joseph W Haus	Rensselaer Polytechnic Institute, Troy, NY Mode-Locked Laser Models and Simulations	RL/OCPA	4 - 12

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
DR James P LeBlanc	New Mexico State University, Las Cruces, NM Multichannel Autoregressive Modeling & Spectral Estimation Methods for Airborne Radar Environment	RL/OCSS	4 - 13
DR David J McLaughlin	Northeastern University, Boston, MA A Review of Microwave Terrain Clutter Measurements at Bistatic	RL/ERCS	4 - 14
DR Hrushikesh N Mhaskar	Cal State Univ, Los Angeles, Los Angeles, Ca Neural Beam Steering & Direction Finding	RL/ERAA	4 - 15
DR Ronald W Noel	Rensselaer Polytechnic Institute, Troy, NY A Low Dimensional Categorization Technique for C Source Code	RL/C3CA	4 - 16
DR Jeffrey B Norman	Vassar College, Poughkeepsie, NY Frequency Response of Semiconductor Photorefractive Matls: ZnTe:Mn:V,GaAs:Cr,&CdMnTe:V	RL/OCPA	4 - 17
DR Glenn E Prescott	University of Kansas Center for Research, Lawrence, KS Rapid Prototyping of Software Radio Sys Using Field Programmable Gate Arrays & DSP Microprocessors	RL/C3BB	4 - 18
DR Mark R Purtill	Texas A&M Univ-Kingsville, Kingsville, TX A Network Flow Heuristic for Graph Mapping	RL/C3CB	4 - 19
DR Mysore R Rao	Rochester Inst. Of Technology, Rochester, NY Detection of Concealed Objects in Images: Investigation into Wavelet Transform Based Object Isolation Techniques	RL/OCSM	4 - 20
DR Scott E Spetka	SUNY of Tech Utica, Utica, NY Integrating a Multimedia Database & WWW Indexing Tools	RL/IRD	4 - 21
DR Gang Sun	University of Massachusetts-Boston, Boston, MA Confined Optical Phonon Modes in Si/ZnS Superlattices	RL/EROC	4 - 22

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Mohammad S Alam	Purdue University, Fort Wayne, IN Fast Infrared Image Registration and High Resolution Reconstruction for Real Time Applications	WL/AAJT	5 - 1
DR Dominick Andrisani II	Purdue University, West Lafayette, IN A Fast Fourier Transform Analysis of Pilot Induced Oscillations	WL/FIGC	5 - 2
DR Pnina Ari-Gur	Western Michigan University, Kalamazoo, MI Texture and Microstructure of Hot Rolled Ti-6Al-4V	WL/MLLN	5 - 3
DR James D Baldwin	University of Oklahoma, Norman, OK Statistical Analysis of Fatigue Crack Growth Rate Data for 7075-T6 Aluminum Damaged by Prior Corrosion	WL/FIB	5 - 4
DR Armando R Barreto	Florida International Univ, Miami, FL Deconvolution of The Space-Time Radar Spectrum	WL/AAMR	5 - 5
MR Larry A Beardsley	Univ of Texas at Austin, Austin, TX The Use of Wavelets and Neural Networks in Data Compression, Data Fusion and Their Effects on Target Identification	WL/MNGA	5 - 6
DR Raj K Bhatnagar	University of Cincinnati, Cincinnati, OH Variable Width Template Construction for ATR with HRR Data	WL/AACR	5 - 7
DR Alley C Butler	University of Cincinnati, Cincinnati, OH Importance of Current Crowding and Self-Heating in a CdS/LaS Cold Cathode	WL/MLIM	5 - 9
DR Reaz A Chaudhuri	University of Utah, Salt Lake City, UT A Novel Compatibility/Equilibrium Based Iterative Post-Processing Approach for Axisymmetric Brittle	WL/MLBM	5 - 11
DR Julian Cheung	New York Inst. Of Technology, New York, NY New Techniques for Non-Cooperative Target Identification	WL/AACT	5 - 12
DR Milton Cone	Embry-Riddle Aeronautical University, Prescott, AZ Of Match Maker and Metrics	WL/AACF	5 - 13
DR Robert R Criss	Randolph-Macon Woman's College, Lynchburg, VA Optical Studies of Two Novel Electro-Explosive Devices	WL/MNMF	5 - 14

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Robert J DeAngelis	Univ of Nebraska - Lincoln, Lincoln, NE Granin Size Effects in the Determination of X-Ray Pole figures and Orientation Distribution Function	WL/MNM	5 - 15
DR Yujie J Ding	Bowling Green State University, Bowling Green, OH Investigation of Photoluminescence Intensity Saturation and Decay, and Nonlinear Optical Devices in Semiconductor Structures	WL/AADP	5 - 16
DR Gregory S Elliott	Rutgers State Univ of New Jersey, Piscataway, NJ Laser Based Diagnostic Techniques for Combustion and Compressible Flows	WL/POPT	5 - 17
DR Altan M Ferendeci	University of Cincinnati, Cincinnati, OH Vertical 3-D Interconnects for Multichip Modules	WL/AADI	5 - 18
DR Dennis R Flentge	Cedarville College, Cedarville, OH Kinetic Studies of the Thermal Decomposition of Demnum and X-1P Using the System for Thermal Diagnostic Studies (STDS)	WL/POSL	5 - 19
DR Himansu M Gajiwala	Tuskegee University, Tuskegee, AL Novel Approach for the Compressive Strength Improvement of Rigid Rod Polymers	WL/MLBP	5 - 20
DR Allen G Greenwood	Mississippi State University, Mississippi State, MS A Framework for Manufacturing-Oriented, Design-Directed Cost Estimation	WL/MTI	5 - 21
DR Rita A Gregory	Georgia Inst of Technology, Atlanta, GA Affects of Int'l Quality Standards on Bare Base Waste Disposal Alternatives	WL/FIVC	5 - 22
DR Michael A Grinfeld	Rutgers University, Piscataway, Piscataway, NJ Mismatch Stresses, Lamellar Microstructure & Mech	WL/MLLM	5 - 23
DR Awatef A Hamed	University of Cincinnati, Cincinnati, OH Inlet Distortion Test Considerations for High Cycle Fatigue in Gas Turbine Engines	WL/FIM	5 - 24
DR Stewart M Harris	SUNY Stony Brook, Stony Brook, NY Compositional Modulation During Epitaxial Growth of Some III-V Heterostructures	WL/MLPO	5 - 25
DR Larry S Helmick	Cedarville College, Cedarville, OH Effect of Humidity on Wear of M-50 Steel with a Krytox Lubricant	WL/MLBT	5 - 26
DR Kenneth L Hensley	University of Oklahoma, Norman, OK Hyperbaric Oxygen Effects on the Postischemic Brain	MED/SGP	5 - 27

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Iqbal Husain	University of Akron, Akron, OH Fault Analysis & Excitation Requirements for Switched Reluctance Starter-Generators	WL/POOC	5 - 28
DR David W Johnson	University of Dayton, Dayton, OH In Situ Formation of Standards for the Determination of Wear Metals in Perfluoropolyalkylether Lubricating Oils	WL/MLBT	5 - 29
DR Marian K Kazimierczuk	Wright State University, Dayton, OH Aircraft Super Capacitor Back-Up System	WL/POOC	5 - 30
DR Edward T Knobbe	Oklahoma State University, Stillwater, OK Corrosion Resistant Sol-Gel Coatings for Aircraft Aluminum Alloys	WL/MLBT	5 - 31
DR Michael C Larson	Tulane University, New Orleans, LA Cracks at Interfaces in Brittle Matrix Composites	WL/MLLM	5 - 32
DR Douglas A Lawrence	Ohio University, Athens, OH Analysis & Design of Gain Scheduled Missile Autopilots	WL/MNAG	5 - 33
DR Junghsen Lieh	Wright State University, Dayton, OH Determination of 3D Deformations, Forces and Moments of Aircraft Tires with a Synchronized Optical and Analog System	WL/FIVM	5 - 34
DR Chun-Shin Lin	Univ of Missouri - Columbia, Columbia, MO Neural Network Technology for Pilot-Vehicle Interface & Decision Aids	WL/FIGP	5 - 35
DR Zongli Lin	SUNY Stony Brook, Stony Brook, NY Control of Linear Sys with Saturating Actuators with Applications to Flight Control Systems	WL/FI	5 - 36
DR Kuo-Chi Lin	University of Central Florida, Orlando, FL Study on Dead Reckoning Translation in High Level Architecture	WL/AASE	5 - 37
DR James S Marsh	University of West Florida, Pensacola, FL A Conceptual Model for Holographic Reconstruction & Minimizing Aberrations During Reconstruction	WL/MNSI	5 - 38
DR Paul Marshall	University of North Texas, Denton, TX Computational Studies of the Reactions of CH3I With H and OH	WL/MLBT	5 - 39

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Hui Meng	Kansas State University, Manhattan, KS Investigation of Holographic PIV and Holographic Visualization techniques for Fluid Flows and Flames	WL/POSC	5 - 40
DR Douglas J Miller	Cedarville College, Cedarville, OH Band Gap Calculations on Oligomers with an All-Carbon Backbone	WL/MLBP	5 - 41
DR Ravi K Nadella	Wilberforce University, Wilberforce, OH Hydrogen & Helium Ion Implantations for Obtaining High-Resistance Layers in N-Type 4H Silicon Carbide	WL/MLPO	5 - 42
DR Krishna Naishadham	Wright State University, Dayton, OH Hydrogen & Helium Ion Implantations for Obtaining High-Resistance Layers in N-Type 4H Silicon	WL/MLPO	5 - 43
DR Timothy S Newman	Univ of Alabama at Huntsville, Huntsville, All A Summer Faculty Project for Anatomical Feature Extraction for Registration of Multiple Modalities of Brain MR	WL/AACR	5 - 44
DR Mohammed Y Niamat	University of Toledo, Toledo, OH FPGA Implementation of the Xpatch Ray Tracer	WL/AAST	5 - 45
DR James L Noyes	Wittenberg University, Springfield, OH The Development of New Learning Algorithms	WL/AACF	5 - 46
DR Anthony C Okafor	University of Missouri - Rolla, Rolla, MO Assessment of Developments in Machine Tool Technology	WL/MTI	5 - 47
DR Paul D Orkwis	University of Cincinnati, Cincinnati, OH Assessing the Suitability of the CFD++ Algorithm for Advanced Propulsion Concept simulations	WL/POPS	5 - 48
Dr Robert P Penno	University of Dayton, Dayton, OH Grating Lobes in Antenna Arrays	WL/AAMP	5 - 49
DR George A Petersson	Wesleyan University, Middletown, CT Absolute Rates for Chemical Reactions	WL/MLBT	5 - 50
DR Mohamed N Rahaman	University of Missouri - Rolla, Rolla, MO Effect of Solid Solution Additives on the Densification & Creep of Granular Ceramics	WL/MLLN	5 - 51

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Martin Schwartz	University of North Texas, Denton, TX AB Initio Modeling of the Enthalpies of Formation of Fluorocarbons	WL/MLBT	5 - 52
DR Thomas E Skinner	Wright State University, Dayton, OH A Method for Studying Changes in Tissue Energetics Resulting from Hyperbaric Oxygen Therapy	MED/SGP	5 - 53
DR Marek Skowronski	Carnegie Melon University, Pittsburgh, PA Investigation of Structural Defects in 4H-SiC Wafers	WL/MLPO	5 - 54
DR Grant D Smith	Univ of Missouri - Columbia, Columbia, MO Theoretical Investigation of Phthalocyanin Dimers	WL/MLPJ	5 - 55
DR James A Snide	University of Dayton, Dayton, OH Aging Aircraft: Preliminary Investigation of Various Materials and Process Issues	WL/MLLP	5 - 56
DR Yong D Song	North Carolina A & T State University, Greensboro, NC Memory-Base Control Methodology with Application to EMRAAT Missile	WL/MNAG	5 - 57
DR Raghavan Srinivasan	Wright State University, Dayton, OH Microstructural Development During Hot Deformation	WL/MLIM	5 - 58
DR Janusz A Starzyk	Ohio University, Athens, OH Feature Selection for ATR Neural Network Approach	WL/AACA	5 - 59
DR Alfred G Striz	University of Oklahoma, Norman, OK On Multiobjective Function Optimization in Engineering Design	WL/FIB	5 - 60
DR Barney E Taylor	Miami Univ - Hamilton, Hamilton, OH Optical and Electro-Optical Studies of Polymers	WL/MLBP	5 - 61
DR Joseph W Tedesco	Auburn University, Auburn, AL Effects of Airblast Characteristics on Structural Response	WL/MNSA	5 - 62
DR Scott K Thomas	Wright State University, Dayton, OH The Effects of Curvature on the Performance of a Spirally-Grooved Copper-Ethanol Heat Pipe	WL/POOS	5 - 63
DR James P Thomas	University of Notre Dame, Notre Dame, IN Subcritical Crack Growth of Ti-6Al-4V Under Ripple Loading Conditions	WL/MLLN	5 - 64
DR Karen A Tomko	Wright State University, Dayton, OH Grid Level Parallelization of an Implicit Solution of the 3D Navier-Stokes Equations	WL/FIM	5 - 65

SRP Final Report Table of Contents

Author	University/Institution	Arnold Engineering Development Center	Vol-Page
Report Title	Directorate		
DR Saad A Ahmed	King Fahd Univ of Petroleum & Minerals, Saudi, Arabia	AEDC	6 - 1
	Turbulence Statistics & Energy Budget of a Turbulent Shear Layer		
DR Csaba A Biegl	Vanderbilt University, Nashville, TN	AEDC	6 - 2
	Turbine Engine Blade Vibration Analysis System		
DR Frank G Collins	Tennessee Univ Space Institute, Tullahoma, TN	AEDC	6 - 3
	Laser Vapor Screen Flow Visualization Technique		
DR Randolph S Peterson	The University of the South, Sewanee, TN	AEDC	6 - 4
DR Robert L Roach	Tennessee Univ Space Institute, Tullahoma, TN	AEDC	6 - 5
	A Process for Setting Up Computation of Swirling Flows in the AEDC H-3 Heater		

SRP Final Report Table of Contents

Author	University/Institution Report Title	U.S. Air Force Academy Directorate	Vol-Page
DR Ryoichi Kawai	Univ of Alabama at Birmingham, Birmingham, AL A Massively Parallel Ab Initio Molecular Dynamics Simulation of Polymers & Molten Salts	USAF	6 - 6

SRP Final Report Table of Contents

Author	University/Institution Report Title	Air Logistic Centers Directorate	Vol-Page
DR Sandra A Ashford	University of Detroit Mercy, Detroit, MI Evaluation of Current Jet Engine Performance Parameters Archive, Retrieval and Diagnostic System	OCALC	6 - 7
MR Jeffrey M Bigelow	Oklahoma Christian Univ of Science & Art, Oklahoma City, OK Enhancing Tinker's Raster-to-Vector Capabilities	OCALC	6 - 8
DR K M George	Oklahoma State University, Stillwater, OK A Computer Model for Sustainability Ranking	OCALC	6 - 9
DR Jagath J Kaluarachichi	Utah State University, Logan, UT Optimal Groundwater Management Using Genetic Algorithm	OCALC	6 - 10

INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from \$20,000 to \$25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1996 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After

completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education*, *Winds of Change*, and *IEEE Spectrum*. Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1996 are shown in this table.

1996 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	572	188	39
(HBCU/MI)	(119)	(27)	(5)
GSRP	235	109	7
(HBCU/MI)	(18)	(7)	(1)
HSAP	474	138	8
TOTAL	1281	435	54

4. SITE VISITS

During June and July of 1996, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MI)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCU/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant

increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1996 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1996 SRP selected participants are shown here.

1996 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	85	123
USAF Laboratory Funds	37	19	15
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	10	5	0
TOTAL	188	109	138

SFRP - 150 were selected, but nine canceled too late to be replaced.

GSRP - 90 were selected, but five canceled too late to be replaced (10 allocations for the ALCs were withheld by AFOSR.)

HSAP - 125 were selected, but two canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1996 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996
Faculty Members	\$690	\$718	\$740	\$740	\$740	\$770
Graduate Student (Master's Degree)	\$425	\$442	\$455	\$455	\$455	\$470
Graduate Student (Bachelor's Degree)	\$365	\$380	\$391	\$391	\$391	\$400
High School Student (First Year)	\$200	\$200	\$200	\$200	\$200	\$200
High School Student (Subsequent Years)	\$240	\$240	\$240	\$240	\$240	\$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of \$50/day for faculty and \$40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and

GSRP associates, 65 % (194 out of 297) claimed travel reimbursements at an average round-trip cost of \$780.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-five percent (85 out of 188) of faculty associates took orientation trips at an average cost of \$444. By contrast, in 1993, 58 % of SFRP associates took orientation visits at an average cost of \$685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

In 1996, RDL implemented direct deposit as a payment option for SFRP and GSRP associates. There were some growing pains. Of the 128 associates who opted for direct deposit, 17 did not check to ensure that their financial institutions could support direct deposit (and they couldn't), and eight associates never did provide RDL with their banks' ABA number (direct deposit bank routing number), so only 103 associates actually participated in the direct deposit program. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1996 REPORT

The complete set of reports for the 1996 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1996 associates, as indicated below:

1996 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. and Puerto Rico. Selectees represent 44 states plus Puerto Rico.

GSRP - Applicants came from 44 states and Puerto Rico. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	188
GSRP	109
HSAP	138
TOTAL	435

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	1	185
Master's	4	48	52
Bachelor's	0	60	60
TOTAL	188	109	297

SFRP Academic Titles	
Assistant Professor	79
Associate Professor	59
Professor	42
Instructor	3
Chairman	0
Visiting Professor	1
Visiting Assoc. Prof.	0
Research Associate	4
TOTAL	188

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28 %	34 %
Colleague familiar with SRP	19 %	16 %
Brochure mailed to institution	23 %	17 %
Contact with Air Force laboratory	17 %	23 %
<i>IEEE Spectrum</i>	2 %	1 %
<i>BIIHE</i>	1 %	1 %
Other source	10 %	8 %
TOTAL	100 %	100 %

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

- A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).
- B. Faculty Associates suggest higher stipends for SFRP associates.
- C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1996 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

Lab	Evals Recv'd	How Many Associates Would You Prefer To Get ? (% Response)											
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
		0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
FJSRL	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
 - a. Time available for evaluation of applications:
 - b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5.
 - a. Benefits of associate's work to laboratory:
 - b. Benefits of associate's work to Air Force:
6.
 - a. Enhancement of research qualifications for LFP and staff:
 - b. Enhancement of research qualifications for SFRP associate:
 - c. Enhancement of research qualifications for GSRP associate:
7.
 - a. Enhancement of knowledge for LFP and staff:
 - b. Enhancement of knowledge for SFRP associate:
 - c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10.
 - a. Your working relationship with SFRP:
 - b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
 b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Responses to above questions

	<i>AEDC</i>	<i>AL</i>	<i>FJSRL</i>	<i>PL</i>	<i>RL</i>	<i>WHMC</i>	<i>WL</i>
<i># Evals Recv'd</i>	0	7	1	14	5	0	46
<i>Question #</i>							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1996 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1996 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13.
 - a. Would you apply again?
 - b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
 - a. At Home:
 - b. With Friend:
 - c. On Local Economy:
 - d. Base Quarters:
17. Value of orientation visit:
 - a. Essential:
 - b. Convenient:
 - c. Not Worth Cost:
 - d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1996 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffin	Hancom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1996 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1996 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

GRATING LOBES IN ANTENNA ARRAYS

Robert P. Penno, Ph.D.
Assistant Professor
Department of Electrical Engineering

University of Dayton
Dayton, Ohio 45469-0226

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air force Base, DC

and

Wright Laboratory

September 1996

GRATING LOBES IN ANTENNA ARRAYS

Robert P. Penno
Assistant Professor
Department of Electrical Engineering
University of Dayton

Abstract

This work explores the problem of grating lobes in an antenna array. First describing the grating lobe, and then the trapped grating lobe, this paper then explores the use of different array feed configurations. Finally, a curvilinear array, as a perturbation of the linear array, is examined for its ability to extinguish the grating lobe. It is seen that certain feed networks can effectively produce nulling of selected grating lobes, while a parabolic curvature of the linear array succeeds only in eliminating the back lobes.

GRATING LOBES IN ANTENNA ARRAYS

Robert P. Penno

Introduction

The primary purpose of an antenna array is to provide enhanced focusing of the transmitted electromagnetic wave. Current technology requires the use of such antenna arrays over a broad band of frequencies, sometimes as much as three times the lowest frequency of the band. When an array is used over such a wide band, its properties vary as the electrical spacing between elements. Thus, antenna elements physically separated by 3 cm display half-wave electrical spacing at 5. GHz, but exhibit an electrical spacing of 1.5 wavelengths at 15 GHz. It is well known that electrical spacing of one wavelength or greater lead to the unwanted appearance of grating lobes. Such lobes represent the transmission of significant portions of power in directions other than the main beam and represent a significant design problem in current, broadband array technology.

Grating Lobes

Given an array of N equivalent point sources as shown in Figure 1, arranged on a line (say along the y -axis), each separated by d meters, the magnitude of the transmitted pattern of the phased array is given by

$$|S_T| = \frac{\sin(N\psi / 2)}{\sin(\psi / 2)},$$

where

$$\psi = \beta d (\sin(\phi) - \sin(\phi_s))$$

It is assumed that ϕ is the angle at which transmitted power is being measured, while ϕ_s is the angle at which the transmitted beam is being aimed, i.e. the scan angle, for which $|S_T| = N$. However, when electrical spacing, d , is one-half wavelength or greater, it is also seen that $|S_T| = N$ at values of $\phi = \phi_g$, where

$$\sin(\phi_g) = \sin(\phi_s) + \frac{m\lambda}{d}$$

where m is an integer and λ is the wavelength. The first grating lobe to appear (as frequency increases) occurs with $m = \pm 1$, and is the primary concern of this section. Figure 2A shows the main lobe prominent in a twelve element array with $.5\lambda$ element spacing and a 30° scan angle. Figure 2B shows the addition of a grating lobe (97.1°) when element spacing has increased to 1.0λ , corresponding to $m = -1$ in the defining equation.

Another concern is the scan angle at which the onset of grating lobes occurs for a given spacing. Figure 3 describes the onset of grating lobes for scan angles from endfire (0°) to broadside (90°). It is observed that as scan angle approaches broadside, larger elemental electrical spacing (and, hence, higher frequencies) is required for grating lobes to occur.

Also of interest is the location of the first grating lobe relative to the location of the main lobe. Figure 4 displays grating lobe location as a function of scan angle, parameterized to elemental electrical spacing. For a scan angle of 40° , for example, at a frequency such that element spacing is $.75 \lambda$, the grating lobe occurs at approximately 120° . For a scan angle of 80° (near broadside) and a spacing of 1.5λ , grating lobes appear at both 32.8° and 119.5° . Thus, as spacing increases, scan angles near broadside will be "pinched" by grating lobes arising from both endfire directions. Where a grating lobe is unavoidable, it is useful to have it located far away from the main beam, or scan angle. Figure 4 displays how higher frequencies of operation, and, hence, increased electrical spacing, produces grating lobes that are closer to the main beam, and more numerous.

Trapped Grating Lobes

Grating lobes are directly dependent upon the spacing of a linear array. In free space, a spacing, d , may not allow grating lobes to exist. However, when that array is embedded in a material characterized by permittivity, ϵ_r , the electrical spacing becomes larger in proportion to the index of refraction, $n (= \sqrt{\epsilon_r})$, possibly large enough to support a grating lobe. The angle at which the grating lobe radiates may be, by Snell's law, beyond the critical angle of the dielectric-free space interface. So, while the main lobe passes through the dielectric-free space interface, the energy in the grating lobe remains trapped in the

antenna layer, influencing the adjacent elements (i.e. element terminal characteristics).

Consider a dielectric-free space layer as shown in Figure 5. Analysis of the grating lobe condition in the material shows that

$$\cos(\phi_g) = \cos(\phi_s) + \frac{m\lambda_m}{d},$$

where λ_m is the wavelength in the material. By applying Snell's law, this becomes

$$\cos(\phi_{g1}) = \cos(\phi_{s1}) + \frac{m\lambda_0}{d}.$$

To establish the conditions for a trapped grating lobe, we let

$$\cos(\phi_s) + \frac{m\lambda_m}{d} < 1 \text{ and } \cos(\phi_{s1}) + \frac{m\lambda_0}{d} > 1.$$

Moreover, for scan angles, ϕ_{s1} , in free space to produce trapped grating lobes,

$$-\frac{m\lambda_0}{d} < \cos(\phi_{s1}) < \sqrt{\epsilon_r} - \frac{m\lambda_0}{d}.$$

As an example, let $d = 1.5 \lambda_0$, $\epsilon_r = 4$. While ϕ_s is 66.4° , ϕ_{s1} is 36.8° . With $m = 1$, it is seen that the grating lobe angle, ϕ_g , in the material is 42.8° , but does not exist in free space. The refractive bending of the dielectric-free space interface has trapped this lobe, converting it to a surface wave.

Such a trapped grating lobe, then, becomes a source of a surface wave, which typically affects the mutual impedance of the array elements. Changes in the mutual impedance lead to perturbations of the input impedance of the array, affecting the amount of radiated power.

Feed Networks

While the number of elements in an antenna array can be exceedingly large, the number of channels, which feed this array is small, typically two or four in number. The ability of the array to perform scanning operations without interference from accompanying grating lobes is limited by the number of channels, but can be improved by modifying the feed networks that connect the elements to the channel(s). What follows is a

comparison of two antenna feed networks as pertains to their ability to filter the accompanying first grating lobe.

Consider an array of $N/2$ elements, spaced $2d$ apart as shown in Figure 6A. If each element is provided with a linear phase shift via a phase shifter, the resulting field pattern due to simple summation of outputs is given by

$$|S| = \frac{\sin(N\Psi/2)}{\sin(\Psi)}$$

where

$$\Psi = \beta d (\sin(\theta_s) - \sin(\theta))$$

$$\theta_s = \text{Scan Angle (relative to normal)}$$

$$\theta = \text{Observation Angle (relative to normal)}$$

This expression, called the Array Factor (or AF), is the benchmark against which other feed networks will be compared.

Two different feed networks are employed here with the intent of extinguishing grating lobes. The first network is shown in Figure 6B. This network uses N elements, spaced d meters apart, with $N/2$ phase shifters, each arranged to feed a pair of adjacent elements. The outputs of these phase shifters are summed, subsequently to be split (e.g. Wilkinson power divider). The outputs of the power divider are then summed after one of the outputs is phase shifted by α_m radians. If the initial phase shifting applied to pair sums is linear (to provide scanning), the output of this array/feed network is given by

$$|S_a| = 2 \cos\left(\frac{\alpha_m}{2}\right) \cos\left(\frac{\beta d \sin(\theta)}{2}\right) \times |S|$$

where α_m is arbitrarily selected to create a null at the first grating lobe. What is apparent is that this feed network separates the N elements into two $N/2$ subarrays, each of which has the field pattern, $|S|$. Thus, the effect of this feed network is to apply the factor, $2\cos(\alpha_m/2)\cos(\beta d \sin(\theta))$,

to the array factor, $|S|$. Since this factor is independent of scan angle, it is also independent of grating lobe locations. Hence, this pattern cannot provide nulls at an arbitrary angle, e.g. a grating lobe.

A variation is described in Figure 6C. Here, the N elements are divided by the network into two, $N/2$ element subarrays, where phase shifters are applied to each element. Each element of subarray one is paired with an adjacent element which comprises subarray two, so that the linear phase shift applied to an element of subarray one is the same as that applied to its complement in subarray two. The sum of these subarrays is phase shifted (relative to each other) by α_m , prior to combining to form the output, given by

$$|S_c| = 2 \cos(\frac{\Psi}{2}) \times |S|.$$

In this network, a grating lobe nulling factor, $2 \cos(\Psi/2)$, by proper selection of the phase shift (α_m), successfully nulls the first grating lobe. Figure 7A shows the pattern, $|S|$, for $d = 1.5\lambda$, a scan angle of 30° , and twelve elements. As expected, the main lobe appears at 30° , but is accompanied by grating lobes at 9.6° , 56.4° , 303.6° , 330° and 350.4° . Figure 7B describes the $|S_c|$, output of the parallel corporate feed network (Figure 6C), where the grating lobes at 56.4° , 9.6° and 330° have been extinguished by the $\cos(\Psi/2)$ factor.

In summary, feed network C is much more effective in its ability to null the first grating lobe, being able to do so wherever that grating lobe appears. It is not assured that this network can null subsequent grating lobes that develop as interelement spacing increases; however, several subsequent grating lobes are nulled as was seen in Figure 7. Feed network A can null a grating lobe only when it appears at certain angles independent of scan angle. The process of combining power outputs of adjacent elements prior to applying the first phase shift, then separating that output before applying the second phase shift (α_m) causes α_m to be independent of the interferometric phase shift, $-\beta d(\sin(\theta))$, that occurs due to the element spacing. When these phase shifts are applied to each element before combining for total output, the total phase shift is $(\alpha_m - \beta d(\sin(\theta)))$. In this case, α_m is no longer independent of $-\beta d(\sin(\theta_s))$, and proper selection of α_m allows nulling of a selected grating lobe.

Curvilinear Arrays

To explore other means of extinguishing unwanted grating lobes, curvilinear arrays were investigated. Consider an N-element array (where N is odd) that lies in the x-y plane, with observation points in the x-y plane (i.e. $\theta = 90^\circ$, $0 \leq \phi \leq 360^\circ$). If the coordinates of the location of the elements is governed by the parabolic equation:

$$y_n = d + bx_n^2,$$

where d and b are arbitrary constants, and the spacing between adjacent elements is maintained at one wavelength, the linearity of the array can be described by the "Latus Rectum" (LA) of the parabola, which is given by $(1/b)$. As b increases, the curvature of the array increases; as b decreases, the array becomes more linear. For the purposes of this investigation, no consideration was given to shading of elements by other elements, i.e. radiation to all observation points was assumed possible. In fact, for azimuth angles of observation near the x-axis ($\phi = 0^\circ$ or 180°), some the elements would block the radiation from those behind them. For a linear array ($b = 0$), such observation angles would coincide with the endfire aspect. In the cases considered here, there is no blockage for near normal observation angles ($\phi \sim 90^\circ$).

Figure 8a describes a near linear array ($b=.00001$, $LA = 10000$), scanned at an azimuth angle of 60° , with a resulting pattern indistinguishable from a 25 element, linear array with one wavelength spacing. The main lobes occur at 60° and 330° , while the grating lobes (of magnitude 25) occur at 120° and 240° . The lobes at 240° and 330° are more correctly considered back lobes, and exist only if the array is in free space. Increasing b to .001 (decrease LA to 1000) produces the curved array shown in Figure 9a, with the field pattern shown in Figure 9b. It is seen that both main lobe ($\phi = 60^\circ$) and grating lobe ($\phi = 120^\circ$) are unattenuated, while the back lobes have shown minor attenuation. This is further displayed in the array of Figure 10a ($b = .01$, $LA = 100$), and accompanying field pattern shown in figure 10b. Here, the main and grating lobes are unattenuated, while the back lobes have been attenuated a significant amount (approximately 12 dB). Moreover, the effects of

curvature in this fashion is to reduce energy in the back lobes, while not affecting the grating lobes.

In summary, grating lobes (and trapped grating lobes) have been described. Two general approaches to minimizing them have been reviewed: special feed networks and curvilinear arrays. While curvature of an array seems an unlikely solution to the problem, application of variations of the corporate feed network cited herein hold great promise for the elimination of selected grating lobes over a broad band of frequencies.

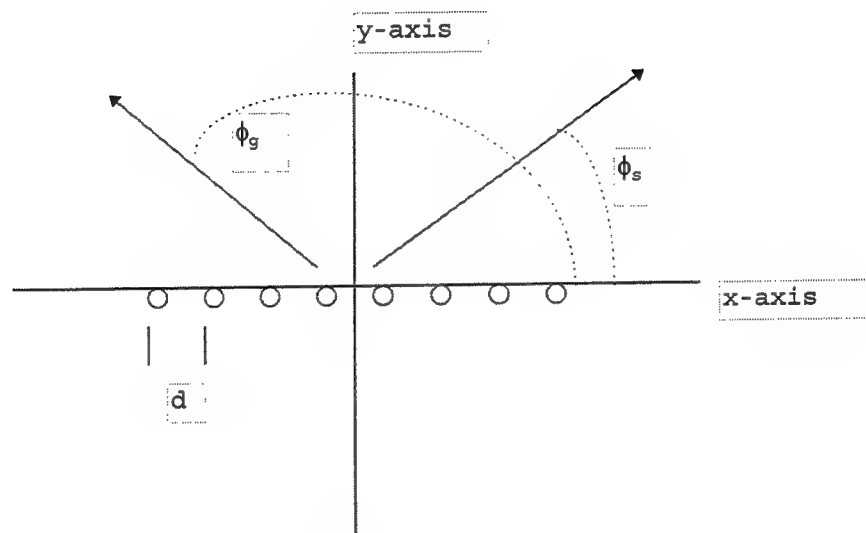


Figure 1: Array Layout

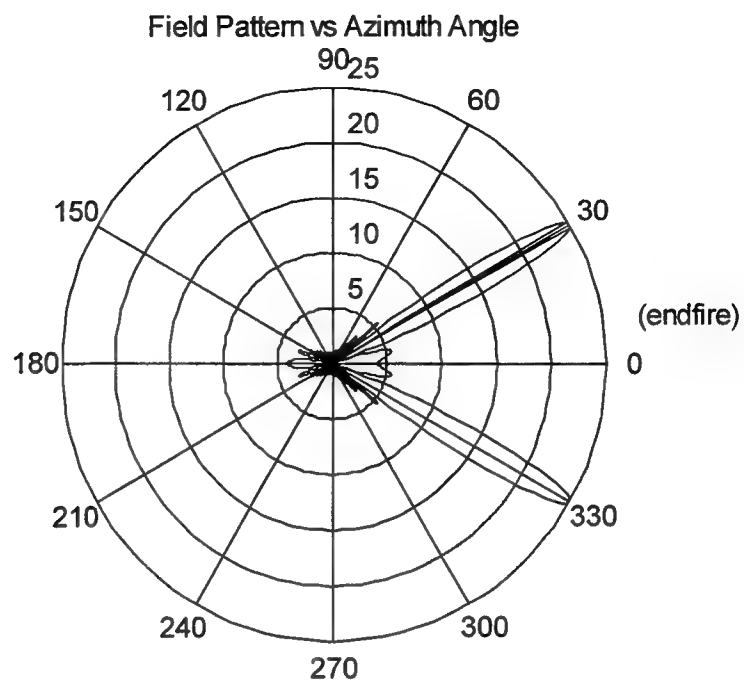


Figure 2A

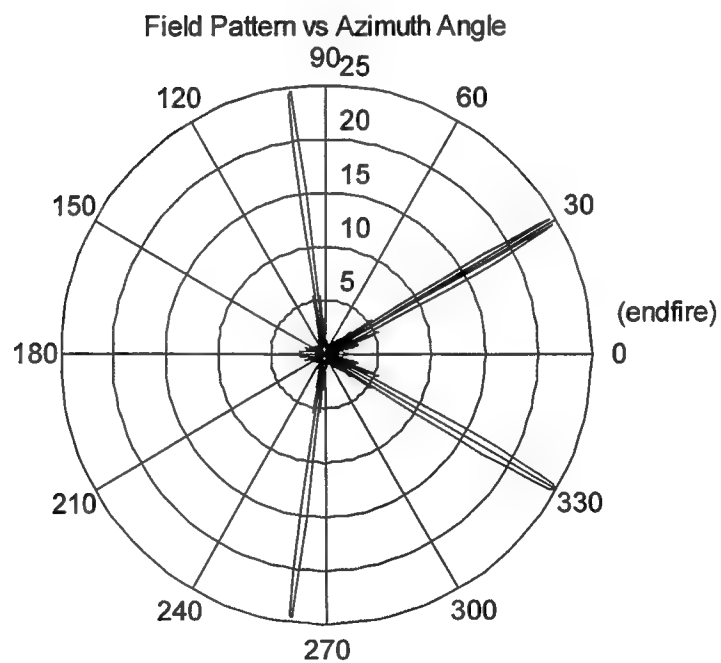


Figure 2B

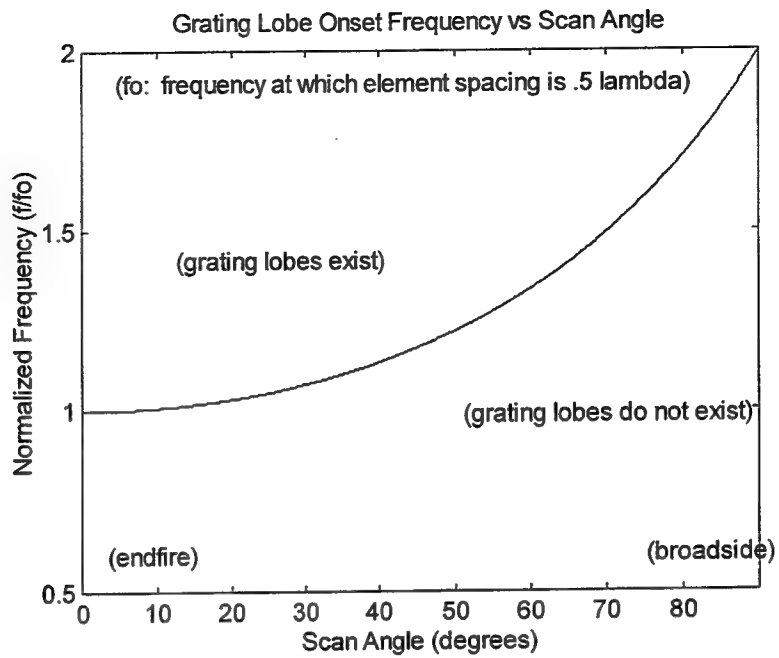


Figure 3

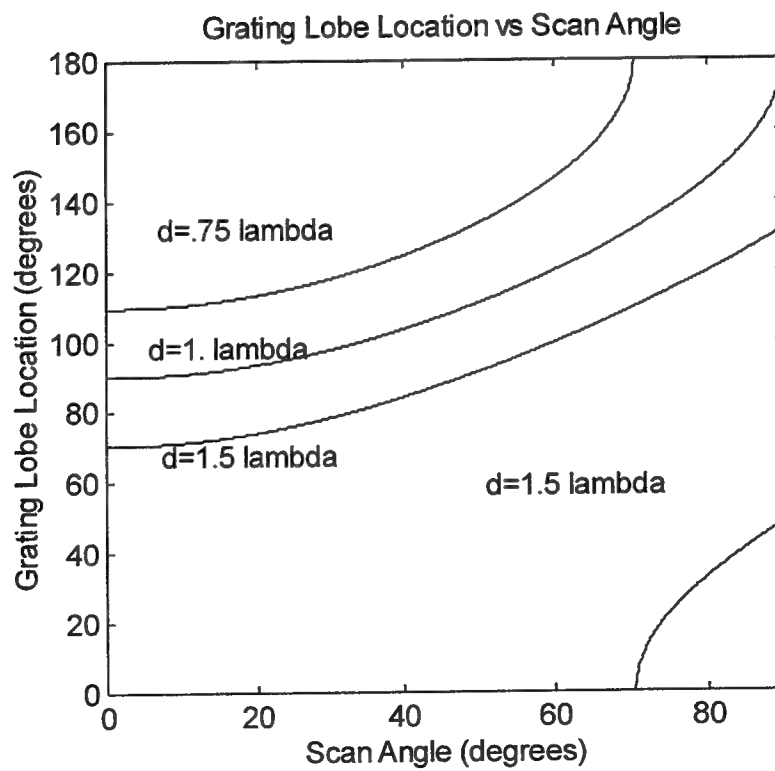


Figure 4

A Dielectric-Free Space Interface

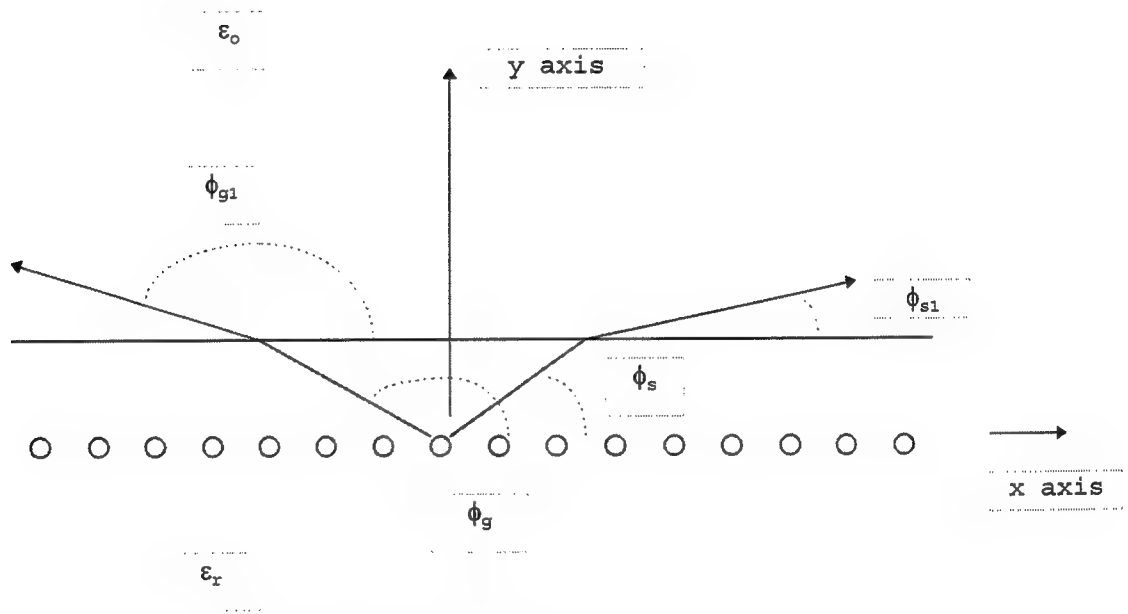


Figure 5

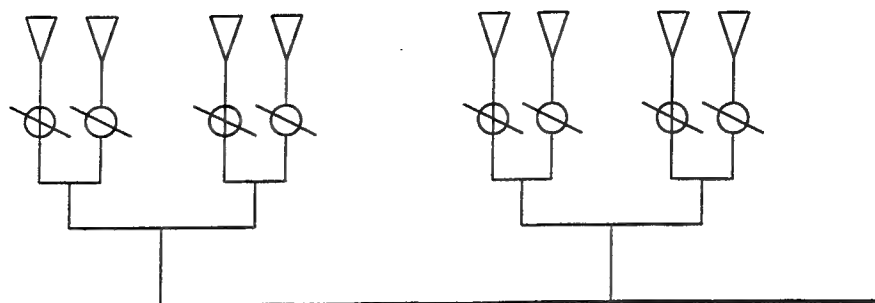


Figure 6 A: Array Factor

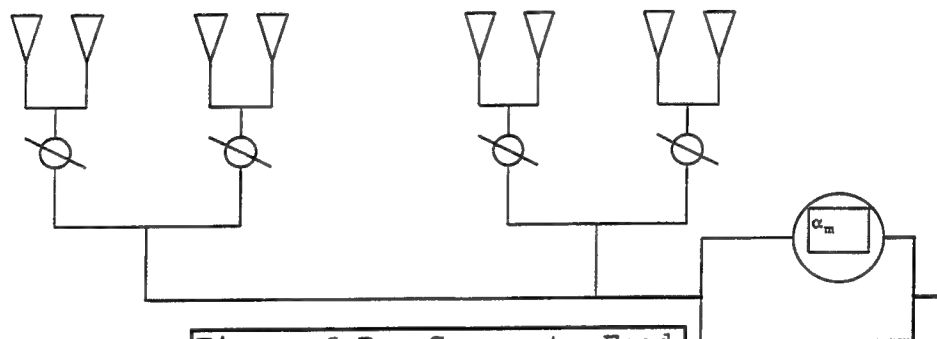


Figure 6 B: Corporate Feed

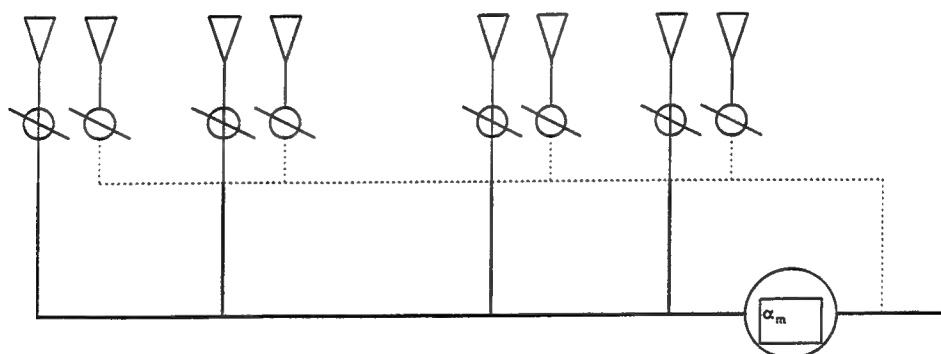
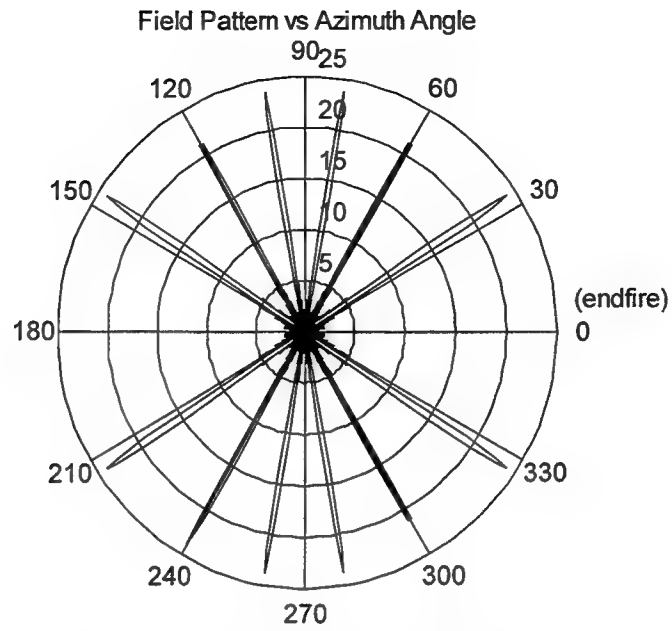
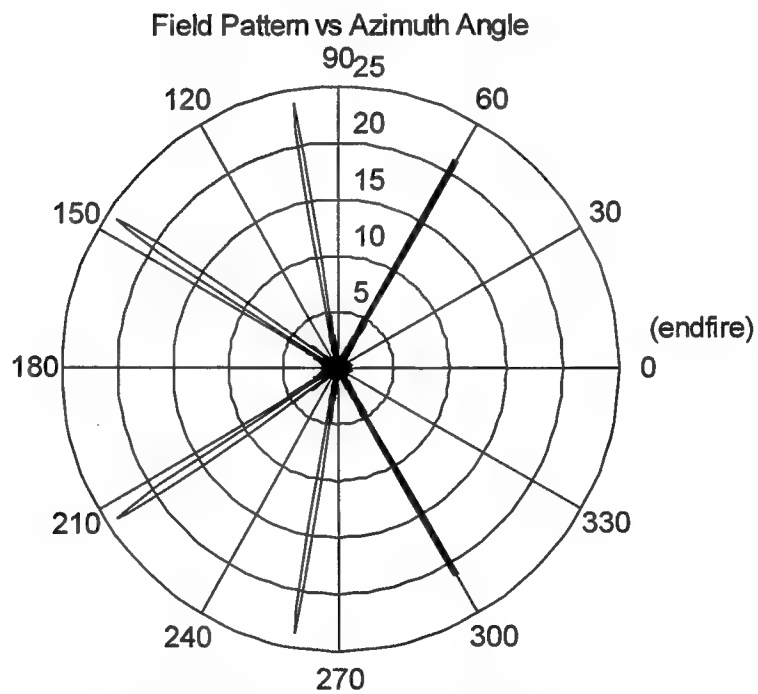


Figure 6 C: Parallel Corporate Feed



$N=24$, Scan Angle = 60° , $d = 1.5 \lambda$

Figure 7A



$N=24$, Scan Angle = 60° , $d = 1.5 \lambda$

Figure 7B

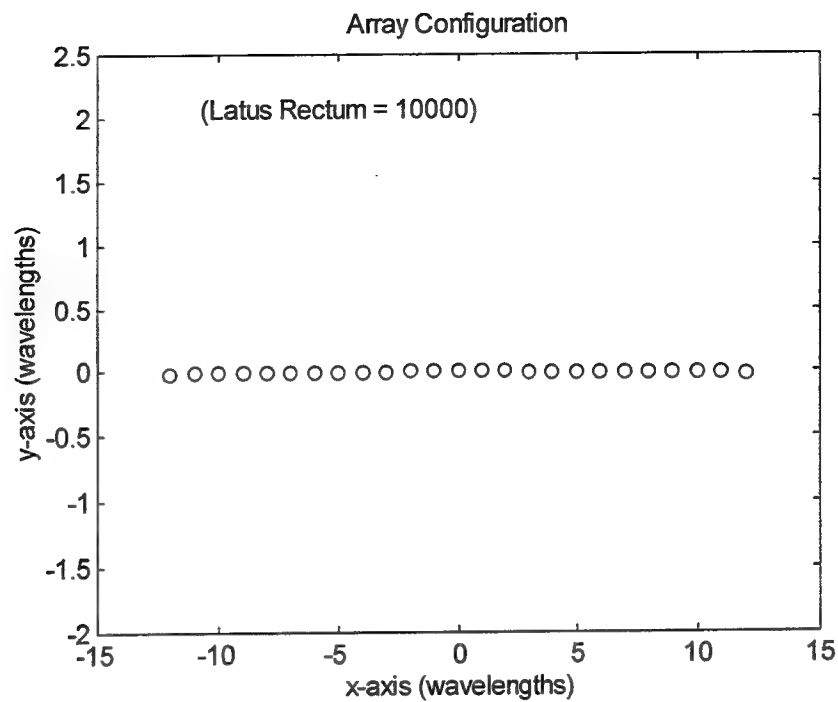


Figure 8A

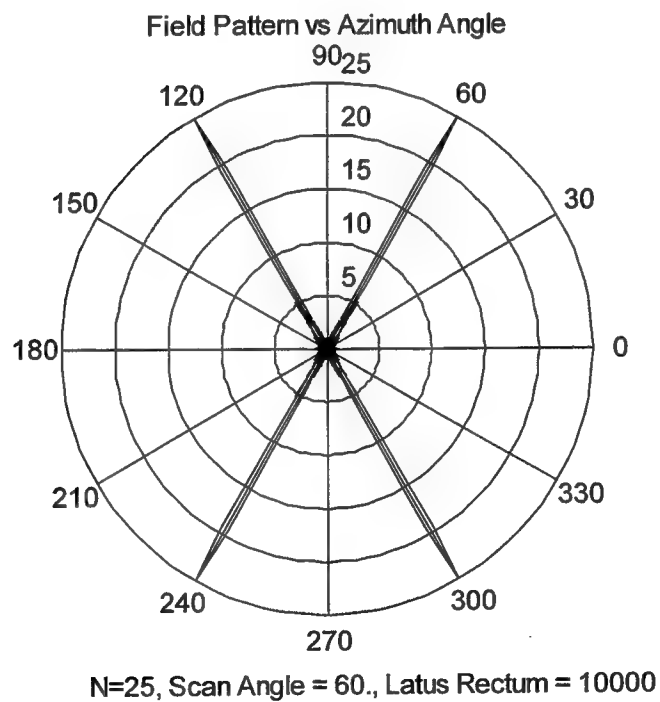


Figure 8B

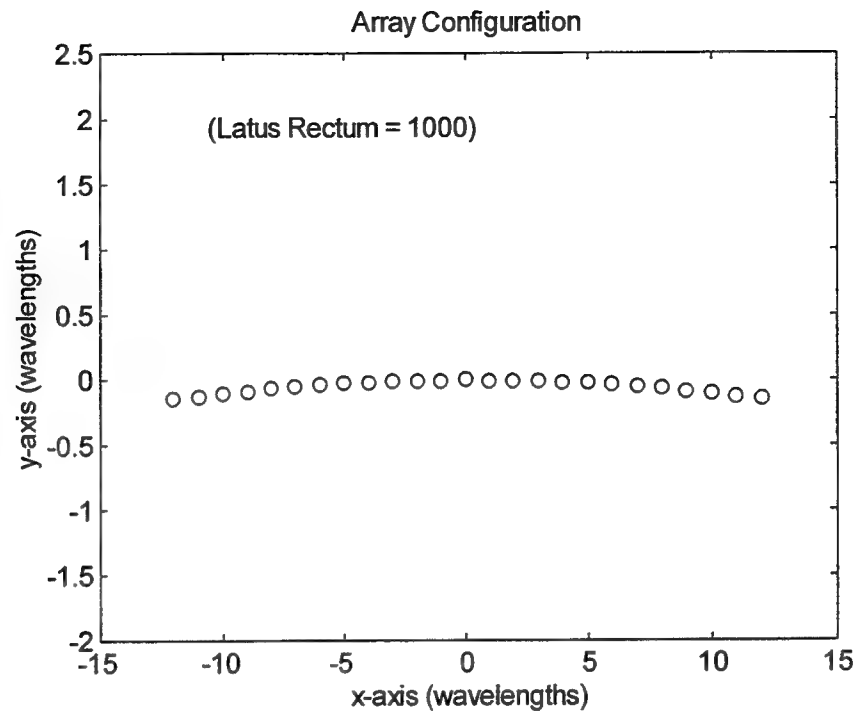
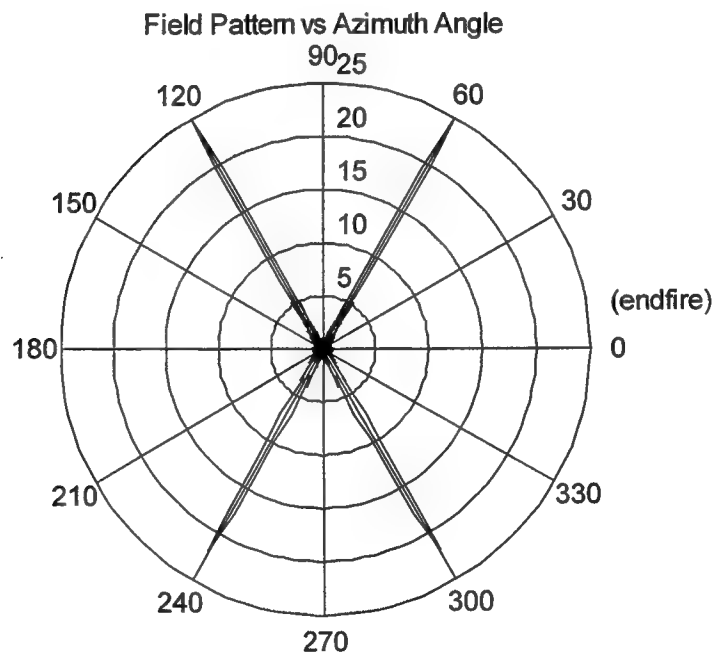


Figure 9A



N=25, Scan Angle = 60., Latus Rectum = 1000

Figure 9B

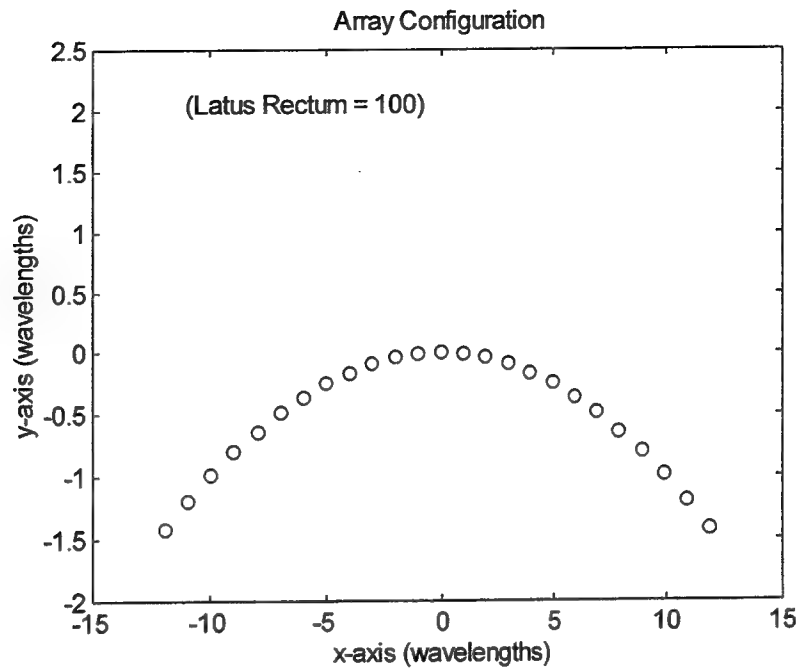
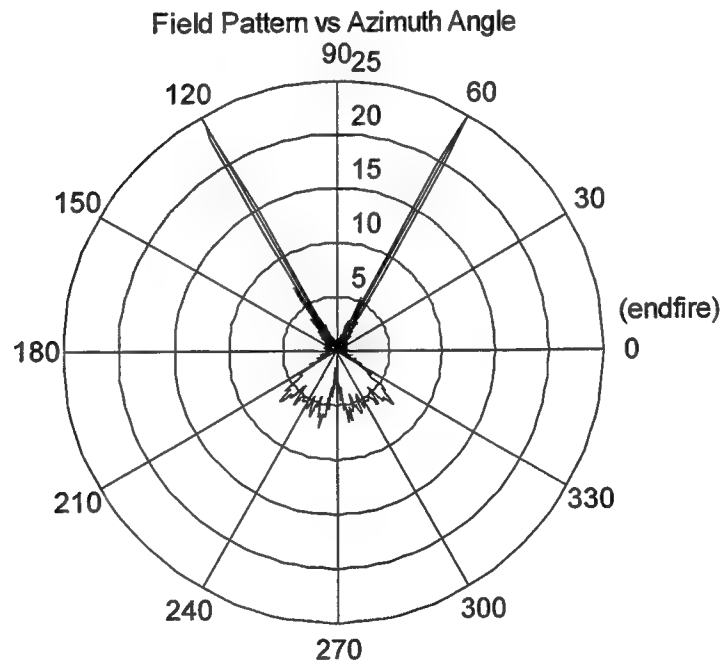


Figure 10A



N=25, Scan Angle = 60., Latus Rectum = 100

Figure 10B

ABSOLUTE RATES FOR CHEMICAL REACTIONS

George A. Petersson
Professor of Chemistry
Chemistry Department

Wesleyan University
Middletown, Connecticut 06459

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

and

Wright Laboratory

August 1996

ABSOLUTE RATES FOR CHEMICAL REACTIONS

G. A. Petersson
Professor of Chemistry
Hall-Atwater Laboratories of Chemistry
Wesleyan University
Middletown, Connecticut 06459

Abstract

There are a number of problems in atmospheric chemistry and flame chemistry that are of considerable importance to the Air Force. A fundamental understanding of the absolute rates of the pertinent chemical reactions will provide the basis for improved solutions to these problems.

A new computational procedure for the characterization of transition states for chemical reactions is proposed and tested. Previous calculations have frequently employed a single point high-level energy calculation at a transition state geometry obtained with a less expensive computational method, $\text{Energy}[\text{Method}(1)]/\text{Geom}[\text{Method}(2)]$. If we instead search the "inexpensive" intrinsic reaction coordinate (IRC) for the maximum of $\text{Energy}[\text{Method}(1)]$ along this reaction path, the resulting "IRCMAX method", $\text{Max}\{\text{Energy}[\text{Method}(1)]\}/\text{IRC}\{\text{Geom}[\text{Method}(2)]\}$, reduces errors in transition state geometries by a factor of four to five, and reduces errors in classical barrier heights by as much as a factor of ten. When applied to the CBS-4, G2(MP2), G2, CBS-Q, and CBS-QCI/APNO model chemistries, the IRCMAX method reduces to the standard model for the reactants and products, and gives RMS errors in the classical barrier heights for ten atom exchange reactions of 1.3, 1.2, 1.0, 0.6, and 0.3 kcal/mol respectively.

The CBS barriers are incorporated into Eyring's absolute rate theory including tunneling through an Eckart potential barrier which includes variations in the zero-point vibrational energies. These calculations reproduce the experimental absolute rates (which vary over eight orders-of magnitude) for five hydrogen transfer reactions with barriers ranging from 1 to 20 kcal/mol at temperatures from 250 K to 2500 K to within 20% RMS relative error.

These methods are now being used to obtain rate constants for important reactions of proposed halon replacements at combustion temperatures. It will soon be possible to experiment with many known and potential agents in many fire scenarios without ever lighting a match or otherwise starting a fire. This is the first step in the construction of a virtual fire laboratory where the computer takes the place of hardware, fuels and suppressants to make possible safe, fast, and effective evaluations of fire suppressants.

ABSOLUTE RATES FOR CHEMICAL REACTIONS

G. A. Petersson

I. Introduction

There are a number of problems in atmospheric chemistry and flame chemistry that are of considerable importance to the Air Force. A fundamental understanding of the absolute rates of the pertinent chemical reactions will provide the basis for improved solutions to these problems.

Halon 1301, trifluoromethyl bromide, has been used as a fire suppressant for everything from computer labs to F-16s for forty years. It is an effective, clean, nontoxic agent that has one major flaw; it is harmful to the earth's protective ozone layer. Because halon was such an ideal agent, research in fire suppression agents had a low priority in Air Force labs and elsewhere until the ozone connection was established. The manufacture of halon has now ceased, and the Air Force is operating on "banked" halon until it runs out, its use is curtailed, or until a satisfactory replacement is found. The last option is preferred, and it is for this reason that the Air Force has major programs to find a replacement.

The Center for Computational Modeling of Nonstructural Materials employs considerable DoD MSRC resources to learn exactly how halon and potential replacement agents work. This program seeks to describe in precise terms all of the important chemical reactions of agents in a flame. Achieving this goal has required the development of improved theoretical methods.

Rate constants describe how fast individual chemical reactions occur, absolutely and in comparison to other reactions. The relative magnitudes permit one to decide which reactions are sufficiently important to include in the overall description of a combustion/suppression mechanism. Because many reactions occur simultaneously, and many more sequentially, it is often impossible to obtain unambiguous individual rate constants in an experiment. This is particularly true at the elevated temperatures typical of combustion reactions. For the relevant reactions which have reliable experimental rate constants, it was found that computationally derived rate constants gave disappointing results. Thus, it was found to be necessary to explore new computational approaches. This required fundamental studies of the theory of chemical reaction kinetics.

The energy of the transition state is of primary importance, and we have developed a new computational method to improve the accuracy of these calculated energies. Also of great importance is the quantum mechanical tunneling, which depends on the width of the potential barrier. This barrier width controls the rates at low temperatures where comparisons with experiment are more easily made. We show that if variations in the zero-point vibrational energy along the reaction path are taken into account, the new kinetics program we have written

quantitatively accounts for tunneling effects. By working with simple reactions, we demonstrate that our new methods give excellent agreement with experimental results for a wide range of temperatures and reaction rates.

These methods are now being used to obtain rate constants for important reactions of proposed halon replacements at combustion temperatures. It will soon be possible to experiment with many known and potential agents in many fire scenarios without ever lighting a match or otherwise starting a fire. This is the first step in the construction of a virtual fire laboratory where the computer takes the place of hardware, fuels and suppressants to make possible safe, fast, and effective evaluations of fire suppressants.

II. Background

The absolute rates of chemical reactions present a formidable challenge to theoretical predictions. An elementary bimolecular chemical reaction:



proceeds at a rate that is proportional to the concentrations of the reactants and the specific rate constant, $k_{rate}(T)$:

$$\frac{dP_1}{dt} = \frac{dP_2}{dt} = -\frac{dR_1}{dt} = -\frac{dR_2}{dt} = k_{rate}(T) R_1 R_2 \quad (2)$$

It is now over one hundred years since Arrhenius published the seminal article¹ on the variation of the rate constant with absolute temperature:

$$k_{rate}(T) = A e^{-\Delta E^\ddagger / RT} \quad (3)$$

The principal difficulty for theoretical predictions lies in the extreme sensitivity of $k_{rate}(T)$ to small errors in the activation energy, ΔE^\ddagger , which can be interpreted as the difference in energy between the "transition state" and the reactants. An error of only 1.4 kcal/mol in ΔE^\ddagger leads to an error of an order-of-magnitude in $k_{rate}(T)$ at room temperature.

More than sixty years ago, Eyring proposed² the use of statistical mechanics to evaluate the preexponential factor, A , in Eq.(3) using partition functions, Q , and the "collision velocity", $k_B T/h$:

$$A(T) = \frac{Q^\ddagger(T)}{Q_{R_1}(T) Q_{R_2}(T)} \cdot \frac{k_B T}{h} \quad (4)$$

thereby accounting for the small temperature variations in A . The evaluation of the translational, vibrational, and rotational partition functions:

$$Q_{\text{translation}} = V (2\pi M k_B T)^{3/2} / h^3 \quad (5)$$

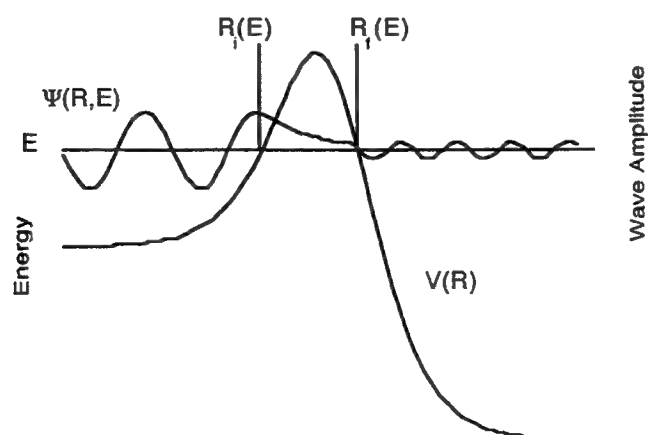
$$Q_{\text{rotation}} = 2\pi (4\pi/\sigma) \prod_j (2\pi I_j k_B T)^{1/2} / h \quad (6)$$

$$Q_{\text{vibration}} = \prod_j (1 - e^{-h\nu_j/k_B T})^{-1} \quad (7)$$

$$Q_{\text{electronic}} = \sum_j e^{-\Delta E_j/k_B T} \quad (8)$$

$$Q_{\text{total}} = Q_{\text{electronic}} Q_{\text{vibration}} Q_{\text{rotation}} Q_{\text{translation}} \quad (9)$$

requires only a knowledge of the mass (for $Q_{\text{translation}}$), geometry (for the moments of inertia, I_j , in Q_{rotation}), vibrational frequencies (ν_j for $Q_{\text{vibration}}$), and ΔE of any low-lying electronic states for the reactants and transition state ($Q_{\text{electronic}}$). Eyring presumed that motion along the path from reactants to products could be treated classically. However, if the reduced mass for this motion is finite, the quantum mechanical wave packet can tunnel through the barrier, $V(R)$, rather than climb over the top:



The probability of reactants with energy, E , tunneling through such a barrier is determined by the initial and final wave amplitudes:

$$K(E) = \frac{\left| \psi \left[R_f(E), E \right] \right|^2}{\left| \psi \left[R_i(E), E \right] \right|^2} \quad (10)$$

Eckart introduced³ a simple mathematical form for the potential energy function, $V(R)$:

$$V(R) = \frac{\Delta E_{\text{reaction}}}{e^{2\pi R/L} + 1} + \frac{\left(\sqrt{\Delta E_{\text{forward}}^\ddagger} + \sqrt{\Delta E_{\text{reverse}}^\ddagger} \right)^2}{e^{2\pi R/L} + 2 + e^{-2\pi R/L}} \quad (11)$$

for which he determined the exact transmission probability, $\kappa(E)$. This transmission probability must then be weighted by the Boltzmann distribution function. The final expression for the rate constant is therefore:

$$k_{\text{rate}}(T) = \frac{Q^\ddagger(T)}{Q_{R_1}(T)Q_{R_2}(T)} \cdot \frac{k_B T}{h} \int_0^\infty \kappa(E) e^{-E/k_B T} dE \quad (12)$$

Truhlar and Kuppermann introduced⁴ the correct definition of the reaction path, R , in mass weighted coordinates, and were the first to recognize the importance of including in $V(R)$, the quantum mechanical zero-point vibrational energy for all normal modes that are orthogonal to R . The challenge we face is to accurately determine all the quantities required to evaluate Eq.(12).

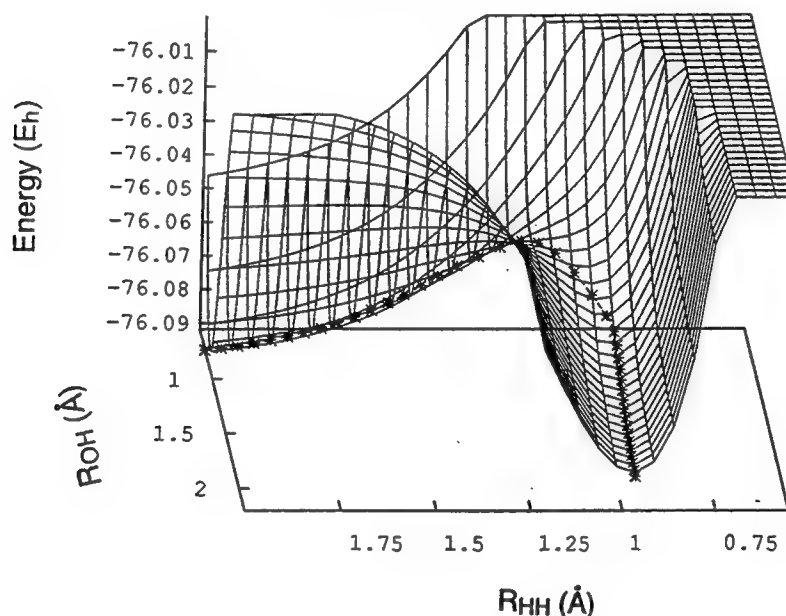
The geometry and energy of a stable molecule can often be measured experimentally to greater accuracy than is currently available computationally. However, transition states cannot generally be isolated for experimental study, and thus are obvious targets for computational studies. The development of analytical gradient and curvature methods⁵⁻¹¹ has made possible the rigorous determination of transition states within a given level of correlation energy and basis set (*e.g.* MP2/6-31G*).¹²

The geometries of stable molecules are relatively easy to predict.¹² Small basis set (*e.g.* 3-21G*) self consistent field (SCF) calculations^{13,14} are generally sufficient to obtain bond lengths to within ± 0.03 Å and bond angles to within $\pm 3^\circ$. However, the prediction of molecular energies to chemical accuracy (*i.e.* ± 2 kcal/mol) requires large basis sets and near full configuration interaction (FCI) levels of correlation energy. The wide gap between the requirements for geometries and energies has spawned a variety of compound models ranging from the simple MP2/6-31G*/UHF/3-21G model to the G2 and CBS series of models.¹⁵⁻²² The compound model notation indicates that the second-order Møller-Plessett (MP2) energy²³ is evaluated with the 6-31G* basis set^{24,25} at the geometry obtained from unrestricted Hartree-Fock (UHF)

calculations^{13,14} with the 3-21G basis set.^{26,27} The G2 and CBS models take the variable basis set approach a step farther, using large basis sets for the SCF energy, medium basis sets for the MP2 energy, and relatively small basis sets for the higher-order correlation energy.²¹ This compound model single point energy is evaluated at a geometry determined at a lower level of theory (*e.g.* CBS-4//UHF/3-21G or G2//MP2/6-31G*). The optimal use of these compound methods for studies of transition states is the subject of this report.

III. The IRCMax Method

The potential energy surface (PES) for a typical bimolecular chemical reaction such as :

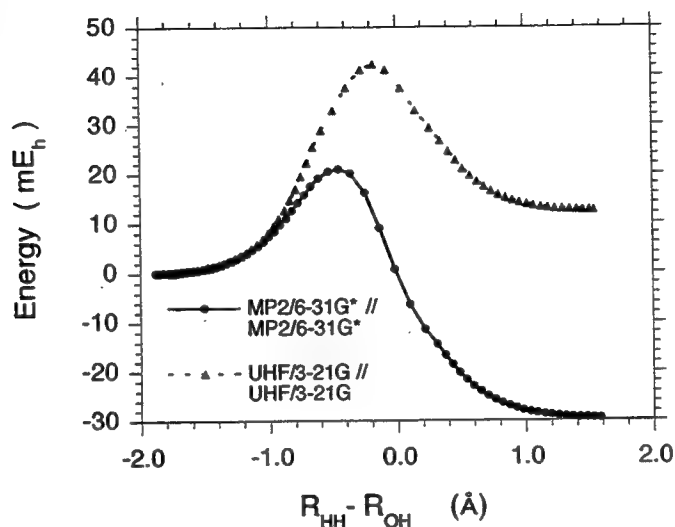


includes valleys (leading to the reactants and products) connected at the transition state (TS), which is a first-order saddle point (*i.e.* a stationary point with exactly one negative force constant). The reaction path or intrinsic reaction coordinate (IRC) is defined²⁸ as the path beginning in the direction of negative curvature away from the TS and following the gradient of the PES to the reactants and products.

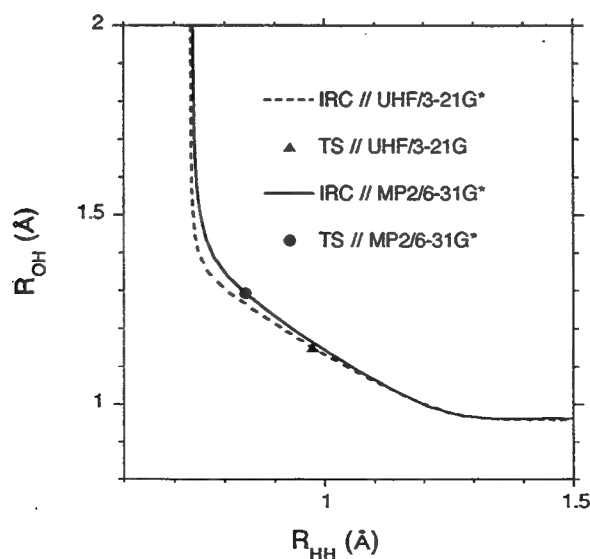
If we move in a direction perpendicular to the reaction path, we find a potential energy curve (or surface) corresponding to a stable reactant or product molecule if we are far from the TS. Even around the TS, the variation of the PES perpendicular to the IRC is very similar to the PES for a stable molecule. Transition States differ from stable molecules in that they possess one negative force constant which defines the reaction coordinate. Calculated energies along the coordinates with positive force constants behave very much like their counterparts in stable molecules. However, the energy changes along the reaction coordinate are much more difficult to

predict. It is the variation of the energy along this coordinate that is very sensitive to (and thus requires the inclusion of) the correlation energy.

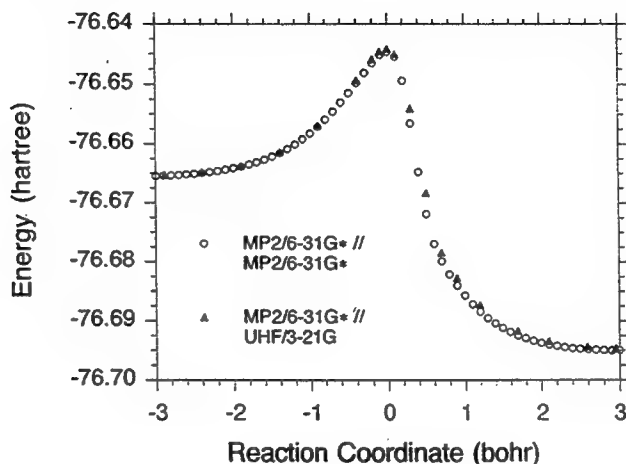
UHF calculations give notoriously bad results for transition states. For example, the UHF/3-21G energy profile for the transfer of a hydrogen atom from H_2 to OH is endothermic rather than exothermic and consequently places the transition state too close to the products:



One might erroneously conclude that such calculations provide no useful information about the reaction path. Fortunately, this is not the case. Although the variation of the energy along the reaction path is very poorly described by the UHF/3-21G method, the variation of the energy perpendicular to the reaction path is very nicely described by the UHF/3-21G method, just as in stable molecules. The UHF/3-21G reaction path very closely approximates the MP2/6-31G* reaction path:



However, the energy variation along this path, and hence the position of the UHF/3-21G transition state, is incorrect. Nevertheless, the UHF/3-21G reaction path passes through (or near) the MP2/6-31G* transition state. Hence, if we calculate the MP2/6-31G* energy along the UHF/3-21G reaction path, we obtain an energy profile that differs trivially from the MP2/6-31G* energy profile along the MP2/6-31G* reaction path:



This is rather remarkable for the $\text{H}_2 + \text{OH}$ reaction, given the very poor UHF/3-21G//UHF/3-21G energy profile.

Based on these observations, we propose the “*IRCM_{ax}*” *transition state method*, in which we select the maximum of the high-level Energy[Method(1)] along the low-level IRC obtained from the Geom[Method(2)] calculations.

The proposed transition state extension of the compound models takes advantage of the enormous improvement in computational speed achieved by using low-level, Geom[Method(2)], IRC calculations. We then perform several single point higher level, Energy[Method(1)], calculations along the Geom[Method(2)] reaction path to locate the Energy[Method(1)] transition state, that is, the maximum of Energy[Method(1)] along the Geom[Method(2)] IRC. Calculations at three points bracketing the transition state are sufficient to permit a parabolic fit to determine the transition state geometry and the activation energy.

Since we determine the maximum of Energy[Method(1)] along a path from reactants to products, the IRCMax method gives a rigorous upper bound to the high-level Method(1) transition state energy. In addition, when applied to the compound CBS and G2 models, the IRCMax method reduces to the normal treatment of bimolecular reactants and products, Energy[Method(1)]//Geom[Method(2)]. Thus the IRCMax method can be viewed as an extension of these compound models to transition states. The results below demonstrate the numerical

superiority of this IRCMax method, $\text{Max}\{\text{Energy}[\text{Method}(1)]//\text{IRC}\{\text{Geom}[\text{Method}(2)]\}$, over conventional $\text{Energy}[\text{Method}(1)]//\text{Geom}[\text{Method}(2)]$ calculations.

IV. Computational Methods

The CBS-4, G2(MP2), G2, CBS-Q, and CBS-QCI/APNO model chemistries have been described in detail elsewhere.¹⁶⁻²¹ Here we shall simply note the RMS errors (2.47, 1.98, 1.54, 1.26, and 0.68 kcal/mol respectively) for the 125 chemical energy differences (*i.e.* D₀'s, IP's, and EA's) of the "G2 test set" to provide a comparison for the activation energies reported in this work.

For the purposes of this report only, the notation for the compound single point energy models (CBS-4, G2, etc.) will refer to their use of varying basis sets for the different orders of perturbation theory, but will not include the geometry optimization methods normally contained in the definition of these methods. Thus the standard CBS-4 and G2 models will be denoted CBS-4//UHF/3-21G* and G2//MP2/6-31G* respectively. The notation, CBS-4, by itself will imply numerical optimization of the geometry with the single point energy method.

Numerical optimization of the geometry of stable molecules and transition state coordinates perpendicular to the reaction path employed parabolic fits on a 0.01 Å grid for bond lengths and 2° for bond angles. Reaction path following with a stepsize of 0.1 bohr (0.0529 Å) was used to determine the IRC's. Parabolic fits to the three highest points determined the position of transition states along the IRC. All calculations employed a modified version of the computer program, Gaussian94.²²

V. Results

The calibration of the accuracy of our calculated transition state geometries and activation energies makes the reasonable assumption that the CBS-QCI/APNO results are the most accurate of the models considered. The initial justification for this assumption rests on comparison with experiment for stable molecules, and comparison with Diedrich and Anderson's accurate numerical results²⁹ for the H₃ transition state (the CBS-QCI/APNO error for the H₃ activation energy is only 0.15 kcal/mol). However, the ultimate vindication comes from the final agreement of our kinetic models with experimental absolute reaction rates.

A. Transition State Geometry

The determination of molecular geometries is the first step in any computational study of molecular properties. Spectroscopic studies of diatomic molecules provide a convenient database³⁰⁻³² of molecular geometries known to within ± 0.0001 Å, and thus can be used to calibrate the accuracy of our computational methods. We have used the diatomic species H₂, CH, NH, OH,

FH, CN, and N₂, and the polyatomic species H₂O, HCN, CH₃, and CH₄ for this geometry calibration. The RMS errors in the UHF/3-21G, MP2/6-31G*, and MP2/6-31G[†] geometries are ±0.012, 0.016, and 0.015 Å respectively. The CBS-4 method is comparable in accuracy (±0.0048 Å) to (but much faster than) the QCISD/6-311G** method (±0.0052 Å). The G2(MP2), G2, and CBS-Q methods are very close in accuracy (±0.0036, 0.0038, and 0.0031 Å respectively). The compound methods (*i.e.* CBS-4, G2(MP2), G2, CBS-Q, and CBS-QCI/APNO) are clearly superior to the simple methods. In fact, the CBS-QCI/APNO method is sufficiently accurate (±0.0013 Å) to be used to calibrate the accuracy of any of the other computational methods.

The calculated geometries of the transition states for six degenerate atom exchange reactions:



where A = H, F, CN, OH, and CH₃, and B = H and F (with A = H) provide an indirect test of the IRCMax concept. Since the position of the transition state along the IRC is determined by symmetry, the accuracy of the calculated geometries depends only on the behavior of the calculated potential energy surface perpendicular to the IRC. Thus, the accuracy of these calculated transition state geometries should be comparable to the accuracy of calculated geometries for stable molecules. In fact, the RMS errors for each of the models (±0.035 Å for UHF/3-21G, ±0.012 Å for MP2/6-31G*, ±0.013 Å for MP2/6-31G[†], ±0.007 Å for QCISD/6-311G**, ±0.009 Å for CBS-4, ±0.0023 Å for G2(MP2), ±0.0033 Å for G2, and ±0.0032 for CBS-Q) is not too much larger than the corresponding error for stable molecules. We conclude that the behavior of these surfaces perpendicular to the IRC's is similar to that of stable molecules, but with somewhat smaller force constants accounting for the modest reduction in the accuracy of calculated geometries. The accuracy of calculated transition state energies should be unaffected, since the reduced force constants make the accuracy of the transition state geometry less critical. The calculated geometries of the transition states for four exothermic hydrogen abstraction reactions:



with R = H and R' = F, OH, or CN, and R = CH₃ with R' = OH provide a direct test of the IRCMax method. We first compare the UHF/3-21G and Max{MP2/6-31G*}/IRC{UHF/3-21G} optimized transition state geometries with the fully optimized MP2/6-31G* geometries, which are the presumed goal of these IRCMax calculations. The relative RMS errors indicate that the IRCMax method reduces geometry errors by a factor of five compared to the UHF/3-21G geometries. The residual relative RMS error (±0.027 Å) is similar to the error in UHF/3-21G geometries for the symmetric transition states for degenerate reactions (±0.035).

We next compare the MP2/6-31G* and Max{QCISD/6-311G**}/IRC{MP2/6-31G*} optimized transition state geometries with the fully optimized QCISD/6-311G** geometries, which

are the goal for these IRCMax calculations. The relative RMS errors indicate that this time the IRCMax method reduces geometry errors by a factor of four compared to the MP2/6-31G* geometries. The residual relative RMS error here (± 0.013 Å) is similar to the error in MP2/6-31G* geometries for the symmetric transition states for degenerate reactions (± 0.012 Å).

Our third test compares the QCISD/6-311G** and Max{CBS-QCI/APNO}/IRC{QCISD/6-311G**} optimized transition state geometries with the fully optimized CBS-QCI/APNO geometries, which are the goal for these IRCMax calculations. The relative RMS errors indicate that this time the IRCMax method reduces geometry errors by a factor of five compared to the QCISD/6-311G** geometries. The residual relative RMS error here (± 0.016 Å) is more than twice as large as the error in QCISD/6-311G** geometries for the symmetric transition states for degenerate reactions (± 0.007 Å). These strongly exothermic reactions have "early" transition states in which the reactants are still quite far apart and only weakly interacting. This makes the transition states less well defined and difficult to locate to better than 0.01 Å. Nevertheless, all three tests of the IRCMax method give substantial improvement over the low level saddle points and demonstrate that the major error in calculated transition state geometries is the location along the IRC as we anticipated.

B. Activation Energy

The enthalpies of reaction, ΔH_{298}^0 , for the four exothermic reactions are known from experiment^{30,33} and thus can be used to calibrate the reliability of each of the computational methods for the energy changes along these reaction paths. The RMS errors are: UHF/3-21G* = 25.4, MP2/6-31G* = 16.3, QCISD/6-311G** = 5.3, CBS-4 = 1.41, G2(MP2) = 1.05, G2 = 1.20, CBS-Q = 0.79, and CBS-QCI/APNO = 0.70 kcal/mol. Given the small size of the current sample set, the agreement with the larger calibration studies^{18,21} is good.

Since the location of the transition state along the reaction path is completely determined by symmetry for the six degenerate reactions, the standard Energy[Method(1)]/Geom[Method(2)] procedure is equivalent to the IRCMax method for these cases. They thus provide a benchmark for the accuracy attainable when the location of the transition state is not in question. In comparison sets of three calculations each, we first calculate the classical barrier height using the simple geometry method, then the higher-level single point energy evaluated at the transition state geometry obtained from the simple geometry method, and finally the energy of the higher-level method evaluated at the transition state geometry obtained from the higher-level method. The relative RMS error is evaluated with respect to the higher-level/higher-level calculation, and the absolute RMS error is evaluated with respect to our best estimate of ΔE_c^\ddagger , the CBS-QCI/APNO model.

The calculated classical barrier heights for the four nondegenerate hydrogen abstraction reactions given in three sets of four calculations each provide a direct test of the IRCMax method. In each set of four calculations, the first row is the calculated classical barrier height using the simple geometry method, the second calculation is the more elaborate single point energy evaluated at the transition state geometry obtained from the simple geometry method, the third calculation is the maximum energy of the more elaborate method evaluated along the IRC obtained from the simple geometry method, and the fourth calculation is the energy of the more elaborate method evaluated at the transition state geometry obtained from the more elaborate method. The fourth calculation is used as the reference to evaluate the relative errors for the first three calculations.

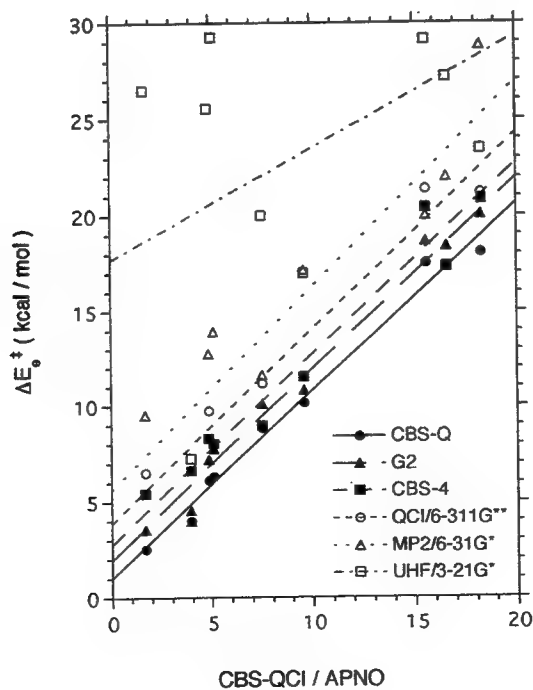
Method/Basis St	RMS Error (kcal/mol)			
	6 Degenerate		4 Nondegenerate	
	Relative	Absolute	Relative	Absolute
UHF/3-21G*	6.37	9.49	13.19	20.87
MP2/6-31G* // UHF/3-21G*	3.38	9.75	6.48	3.70
Max{MP2/6-31G*} // IRC{UHF/3-21G*}			0.59	7.57
MP2/6-31G*	-	7.12	-	7.12
MP2/6-31G*	4.23	7.12	2.93	7.12
QCISD/6-311G**// MP2/6-31G*	1.63	5.13	0.40	4.72
Max{QCISD/6-311G**} // IRC{MP2/6-31G*}			0.14	4.88
QCISD/6-311G**	-	4.07	-	4.79
QCISD/6-311G**	3.93	4.07	4.64	4.64
CBS-QCI/APNO // QCISD/6-311G**	0.09	0.19	0.49	0.49
Max{CBS-QCI/APNO} // IRC{QCISD/6-311G**}			0.24	0.24
CBS-QCI/APNO	-	-	-	-

The first set of four calculations illustrates the behavior of the IRCMax method very clearly. This set compares the UHF/3-21G* method with the MP2/6-31G* energies calculated at UHF/3-21G* and MP2/6-31G* geometries. The UHF barrier heights are all much too high. This is an extreme example of the tendency of calculated barrier heights to converge to the exact value from above. The MP2//UHF barrier heights are all lower than the MP2//MP2 values. The UHF transition state geometry is at the wrong position along the reaction path and thus gives a value for the MP2//UHF transition state energy that is too low. The extreme case is the H-H-F barrier for which the MP2//UHF value is *negative*. In less extreme cases, this negative error can cancel the positive error (from the use of limited correlation/basis sets) to reduce the absolute RMS error. However, the results are erratic with no clear convergence pattern as the basis set and level of

correlation treatment are improved. The IRCMax method gives an upper bound to the MP2 energy of the transition state, but the error from use of the UHF geometry for the reactants can exceed the IRCMax error for the MP2 transition state energy (as for H-H-CN) giving an IRCMax barrier that is below the MP2//MP2 barrier. Nevertheless, the difference between the IRCMax barriers and the MP2//MP2 barriers is minor, and we retain the inherent convergence pattern of calculated barrier heights. The $\text{Max}\{\text{MP2}/6\text{-}31\text{G}^*\}/\text{IRC}\{\text{UHF}/3\text{-}21\text{G}^*\}$ procedure reduces the relative RMS error by a factor-of-ten (from 6.48 to 0.59 kcal/mol).

The calculated classical barrier heights of all ten test reactions are summarized in the table below for the simple UHF, MP2, and QCI methods and for the five recommended IRCMax methods. The eight sets of calculations are arranged in order of decreasing RMS errors. For each reaction, the sequence of improving calculations tends to converge to the exact classical barrier height from above. The CBS-QCI/APNO barrier height for the $\text{H}_2 + \text{H}$ reaction has converged to within 0.15 kcal/mol of the exact barrier obtained by numerical solution of the Schrödinger equation.²⁹ Having verified the accuracy of the CBS-QCI/APNO method in the one case for which the exact barrier height is known, we then use the CBS-QCI/APNO//CBS-QCI/APNO values for the nine remaining reactions to determine the errors in the less computationally demanding methods.

A graph of the barrier heights calculated with the less demanding methods vs the CBS-QCI/APNO barrier heights demonstrates the systematic tendency to overestimate barrier heights.



Linear least-squares fits give straight lines with unit slope and nonzero intercept for all but the simple UHF/3-21G* calculations. Each of these computational methods consistently overestimates all ten barrier heights by about the same error. This suggests a simple empirical correction to the calculated classical barrier heights.

Recall that these compound model chemistries already include two empirical parameters, a one-electron correction and a two-electron correction.¹⁶⁻²¹ In the case of the G2 methods, the one-electron correction [$\text{Const1}(n_{\alpha} + n_{\beta})$] was selected to give the exact energy for the hydrogen atom.¹⁶ The two-electron correction [$\text{Const2}(n_{\beta}=n_{\text{pairs}})$] was selected to give the exact energy for the hydrogen molecule.¹⁶ We therefore propose a third empirical correction selected to give the exact energy for the H_3 transition state [$\text{Const3}(n_{k<0})$]. The calculated classical barrier heights including this empirical correction are given in the table.

Method/Basis Set	H ₂ +F	H ₂ +CN	H ₂ +OH	CH ₄ +OH	H ₂ O +OH	H ₂ +H	HF +F	CH ₄ +CH ₃	HCN +CN	HF +H	Ave Error	RMS Error
UHF/3-21G*	19.11	-0.14	18.18	21.88	12.64	9.61	21.75	19.81	16.09	37.16	5.08	9.40
MP2/6-31G*	1.97	-3.59	5.21	6.40	4.08	9.61	12.45	14.47	21.24	42.93	-1.06	3.07
QCISD/6-311G**	4.82	5.10	8.06	9.39	9.51	9.61	19.67	18.18	19.46	45.91	2.44	2.80
IRCMa												
CBS-4 a	3.47	4.18	6.36	6.09	7.01	9.61	18.48	15.39	18.05	43.33	0.66	1.34
G2(MP2)	2.51	3.44	6.02	6.57	9.11	9.61	17.93	17.22	18.89	42.92	0.89	1.18
G2	2.30	3.31	5.96	6.53	8.87	9.61	17.45	17.17	18.80	42.75	0.74	1.02
CBS-Q b	1.97	3.86	5.57	5.77	8.27	9.61	16.98	16.79	18.38	41.79	0.37	0.62
CBS-QCI/APNO	2.04	3.95	4.95	5.45	7.59	9.61	15.75	16.50	18.26	42.14	0.10	0.17
CBS-QCI/APNO -.15	1.70	3.96	4.87	5.11	7.51	9.61	15.58	16.60	18.28	42.17	-	-

a. The CBS-4 barriers were increased by 3.63 kcal/mol for each CN group.

b. The CBS-Q barriers were increased by 0.43 kcal/mol for each CN group.

All nine methods now give the exact barrier height for H_3 . The average errors are now generally small compared to the RMS errors, demonstrating that the dominant systematic errors have been removed. The correction based on H_3 gives a general improvement for all these atom exchange reactions. An additional empirical correction for the spin contaminated cyano group significantly improved the CBS-4 and CBS-Q barriers. This suggests that a general bond additivity correction (BAC)^{34,35} might also be useful for calculations of activation energies.

Based on our experience²¹ that the CBS-Q errors are half as large as the CBS-4 errors and the CBS-QCI/APNO errors are half as large as the CBS-Q errors, we estimate the uncertainty in the CBS-QCI/APNO barrier heights to be ± 0.3 kcal/mol. This is consistent with the CBS-QCI/APNO overestimation of the H₃ barrier height by 0.15 kcal/mol before the empirical correction was applied. If we use the nonlinear CBS-QCI/APNO optimized geometry for the H₂ + F transition state, our empirically corrected CBS-QCI/APNO barrier height (1.70 ± 0.3 kcal/mol) for this highly exothermic reaction is in agreement with the best CCSD(T) (1.7 ± 0.2 kcal/mol)³⁶ and MRCI calculations (1.53 ± 0.15 kcal/mol).³⁷

All five compound models give chemically accurate barrier heights (*i.e.* RMS error < 2 kcal/mol). The RMS errors are consistently about half as large as those for the same models applied to dissociation energies, ionization potentials, and electron affinities.²¹

C. Vibrational Frequencies

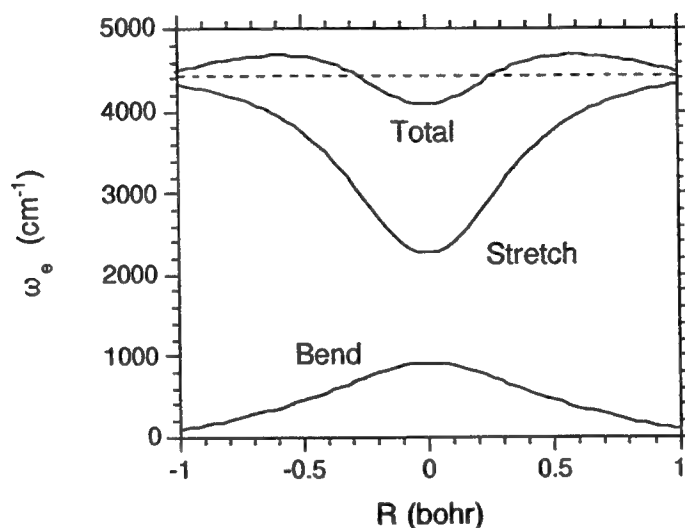
Vibrational energy levels include harmonic components linear in the quantum number, n , and anharmonic corrections:

$$E_n = (n + \frac{1}{2})\hbar\omega_e - (n + \frac{1}{2})^2\hbar\omega_e X_e + \dots \quad (13)$$

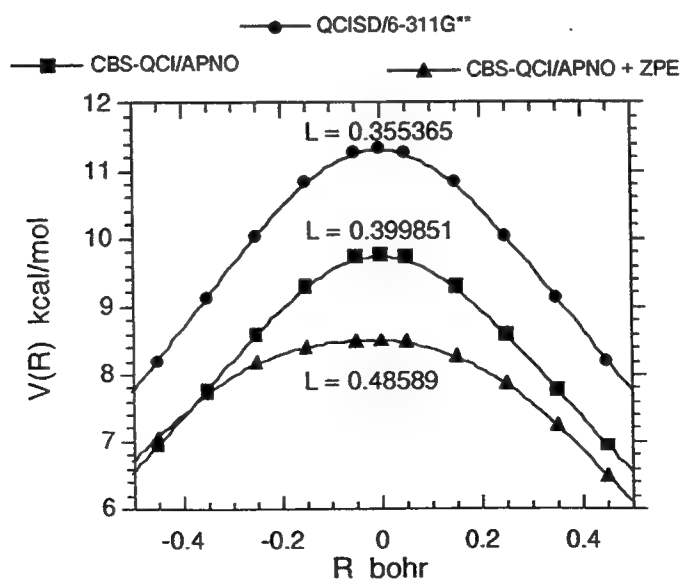
The harmonic term generally dominates the low-lying levels, especially for the stiff normal modes that dominate zero-point energies. For example, the values of ω_e and $\omega_e X_e$ for H₂ are 4401.23 cm⁻¹ and 121.34 cm⁻¹ respectively.³⁰ We therefore employ the harmonic approximation for all calculated frequencies and zero-point energies. The convergence of the harmonic frequencies obtained with the computational methods used for the IRC calculations is illustrated by the results for H₂ and the H₂+H transition state:

Method/Basis Set	H ₂	H ₃		
	Stretch	Stretch	Bend(2)	Rxn Path
UHF/3-21G	4657.1	2054.8	1121.4	2292.2 <i>i</i>
MP2/6-31G*	4533.6	2087.8	1164.6	2325.5 <i>i</i>
QCI/6-311G**	4420.1	2054.2	958.2	1589.1 <i>i</i>
QCI/6-311G(2df,2pd)	4404.1	2051.3	917.1	1530.9 <i>i</i>
Experiment	4401.23			

Note that polarization functions on the central hydrogen of the H--H--H transition state are necessary to obtain an accurate bending frequency. Without them, we would overestimate the ZPE of the transition state by 0.71 kcal/mol. To minimize such errors, we shall employ QCI/6-311G(2df,2pd) frequencies for our initial study of absolute reaction rates. The change in the total ZPE at the H₃ transition state decreases the barrier height, ΔE^\ddagger in Eq.(11), by 0.74 kcal/mol, and thus accelerates the reaction by more than a factor-of-three at room temperature. The variation of the H₃ stretch and bend as we move along the reaction path increases $V(R)$ at $R = \pm 0.5$ bohr to either side of the maximum:



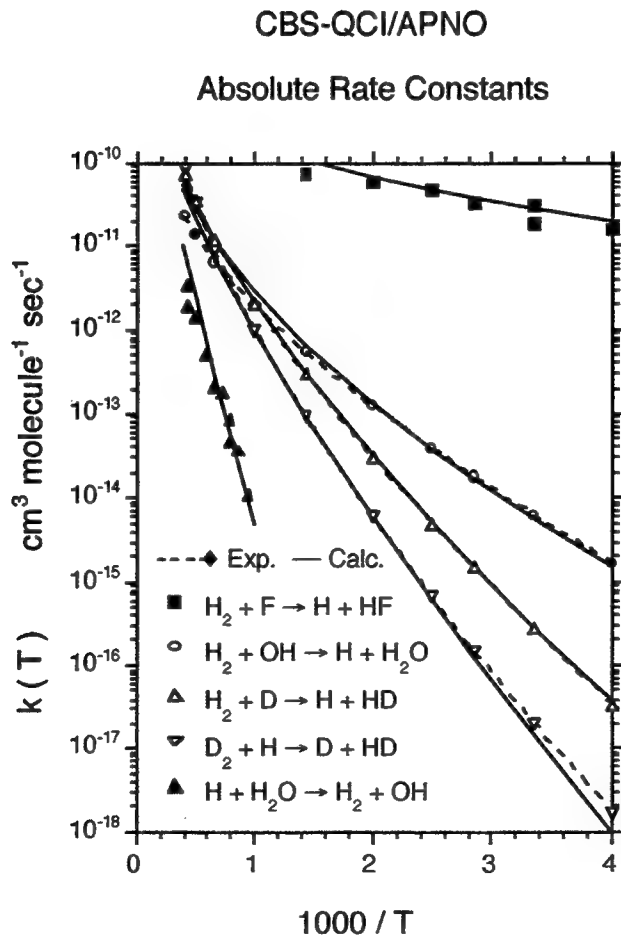
and thus broadens the Eckart barrier, Eq.(11), as illustrated for $\text{H}_2 + \text{D} \rightarrow \text{H} + \text{HD}$:



This increase in the Eckart width parameter, L in Eq.(11), reduces the tunneling through the barrier and hence the calculated reaction rate by a factor-of-ten at low temperatures.

D. Absolute Reaction Rates

We are finally ready to evaluate all of the terms in Eq(12). We have selected six hydrogen abstraction reactions for the initial test of our methodology. The barrier heights for these reactions range from 1 kcal/mol ($\text{H}_2 + \text{F}$) to 20 kcal/mol ($\text{H}_2\text{O} + \text{H}$). We include temperatures from 250 K to 2500 K. The rate constants range from 10^{-18} up to 10^{-10} $\text{cm}^3 / \text{molecule sec}$. All absolute rate constants obtained from our CBS-QCI/APNO model are within the uncertainty of the experiments!³⁸



The dashed curves and open symbols for $\text{H}_2 + \text{OH}$, $\text{H}_2 + \text{D}$, and $\text{D}_2 + \text{H}$ represent the least-squares fits of smooth curves to large experimental data sets in an attempt to reduce the noise level

in the experimental data.³⁸ The close agreement with theory suggests that this attempt was successful.

VI. Conclusions

The problem of predicting absolute rates for gas phase chemical reactions has been solved. However, much remains to be done. We must systematically determine the minimum level of calculation necessary to maintain the accuracy of each part of these calculations. Nevertheless, the use of the IRCMax method with the G2 and CBS-Q computational models is giving immediate improvements in our ability to model flame chemistry.

These methods are now being used to obtain rate constants for important reactions of proposed halon replacements at combustion temperatures. It will soon be possible to experiment with many known and potential agents in many fire scenarios without ever lighting a match or otherwise starting a fire. This is the first step in the construction of a virtual fire laboratory where the computer takes the place of hardware, fuels and suppressants to make possible safe, fast, and effective evaluations of fire suppressants.

References

1. S. Arrhenius, Z. Physikal. Chem. **4**, 226 (1889).
2. H. Eyring, J. Chem. Phys. **3**, 107 (1935).
3. C. Eckart, Phys. Rev. **35**, 1303 (1930).
4. D. G. Truhlar and A. Kuppermann, J. Am. Chem. Soc. **93**, 1840 (1971).
5. C. Peng, and H. B. Schlegel, Israel J. Chem. **33**, 449 (1993).
6. N. C. Handy, and H. F. Schaefer, III, J. Chem. Phys. **81**, 5031 (1984).
7. J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quant. Chem. Symp **13**, 325 (1979).
8. H. B. Schlegel, J. Comp. Chem. **3**, 214 (1982).
9. M. Head-Gordon, and T. Head-Gordon, Chem. Phys. Lett. **220**, 122 (1994).
10. M. J. Frisch, M. Head-Gordon, and J. A. Pople, Chem. Phys. Lett. **166**, 275 (1990).
11. M. J. Frisch, M. Head-Gordon, and J. A. Pople, Chem. Phys. Lett. **166**, 281 (1990).
12. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople *Ab Initio Molecular Orbital Theory*, John Wiley & Sons: New York, 1986.
13. C. C. J. Roothan, Rev. Mod. Phys. **23**, 69 (1951).
14. J. A. Pople, and R. K. Nesbet, J. Chem. Phys. **22**, 571 (1959).
15. J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, J. Chem. Phys. **90**, 5622 (1989).
16. L. A. Curtiss, C. Jones, G. W. Trucks, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **93**, 2537 (1990).

17. L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).
18. L. A. Curtiss, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **98**, 1293 (1993).
19. J. A. Montgomery, Jr., J. W. Ochterski, and G. A. Petersson, J. Chem. Phys. **101**, 5900 (1994).
20. J. W. Ochterski, G. A. Petersson, and K. B. Wiberg, J. Am. Chem. Soc. **117**, 11299 (1995).
21. J. W. Ochterski, G. A. Petersson, and J. A. Montgomery, Jr., J. Chem. Phys. **104**, 2598 (1996).
22. Gaussian94, Revision B.3, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, Jr., K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople; Gaussian, Inc.,: Pittsburgh PA, 1995; pp .
23. C. Møller, and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
24. R. Ditchfield, W. J. Hehre, and J. A. Pople, J. Chem. Phys. **54**, 724 (1971).
25. W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).
26. J. S. Binkley, J. A. Pople, and W. J. Hehre, J. Am. Chem. Soc. **102**, 939 (1980).
27. M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, and W. J. Hehre, J. Am. Chem. Soc. **104**, 2797 (1982).
28. C. Gonzalez, and H. B. Schlegel, J. Phys. Chem. **90**, 2154 (1989).
29. D. L. Diedrich, and J. B. Anderson, Science **258**, 786 (1992).
30. K. P. Huber, and G. Herzberg *Constants of Diatomic Molecules*; Van Nostrand Reinhold: New York, 1979.
31. D. L. Gray, and A. G. Robiette, Molecular Physics **37**, 1901 (1979).
32. E. Hirota, J. Molec. Spect. **77**, 213 (1979).
33. J. Chase M. W. , C. A. Davies, J. Downey J. R. , D. J. Frurip, R. A. McDonald, and A. N. Syverud, J. Phys. Chem. Ref. Data **14**, **Suppl. 1**, (1985).
34. P. Ho, and C. F. Melius, J. Phys. Chem. **94**, 5120 (1990).
35. M. D. Allendorf, and C. F. Melius, J. Phys. Chem. **97**, 72 (1993).
36. G. E. Scuseria, J. Chem. Phys. **95**, 7426 (1991).
37. P. J. Knowles, P. Stark, and H. Werner, Chem. Phys. Letters **185**, 555 (1991).
38. NIST Chemical Kinetics Database, version 6.0.

EFFECT OF SOLID SOLUTION ADDITIVES ON THE
DENSIFICATION AND CREEP OF GRANULAR CERAMICS

M. N. Rahaman
Professor
Department of Ceramic Engineering

University of Missouri-Rolla
222 McNutt Hall
Rolla, MO 65401-0249

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and
Wright Laboratory

August 1996

EFFECT OF SOLID SOLUTION ADDITIVES ON THE DENSIFICATION AND CREEP OF GRANULAR CERAMICS

M. N. Rahaman
Professor
Department of Ceramic Engineering
University of Missouri-Rolla

Abstract

The effect of solid solution additives on the densification and creep of granular ceramics was investigated for a model system consisting of CeO_2 solid solutions with Y_2O_3 as the additive. In the sintering of powder compacts at 1150°C , the densification rate of CeO_2 at a given density decreased significantly with increasing Y^{3+} concentration. The reduction in the densification rate reached a factor of ≈ 100 for a Y^{3+} concentration of 6 atomic percent. Creep of dense cylindrical specimens was investigated at constant strain rates of 10^{-5} and 10^{-4} s^{-1} in air at 1200°C . After compensation for differences in grain size, the creep rate was also found to decrease significantly with increasing Y^{3+} concentration. If the creep rate is assumed to be controlled by a mechanism of grain boundary diffusion, then the magnitude of the decrease is in good agreement with that observed in the sintering experiments. The results strongly indicate that it may be possible to predict changes in the steady-state creep behavior from observed changes in the sintering behavior provided that matter transport occurs by the same mechanism. They also indicate that the solid solution approach may have considerable merit for controlling the creep resistance of rare earth oxides that commonly have a high solubility for many cations.

EFFECT OF SOLID SOLUTION ADDITIVES ON THE DENSIFICATION AND CREEP OF GRANULAR CERAMICS

M. N. Rahaman

1. INTRODUCTION

Densification and creep have long been recognized as fundamental processes in materials. However, over the past 10 to 15 years, their interaction has been shown to have important consequences in several applications. Densification of heterogeneous and constrained systems normally generate incompatibilities in the deformation which require shear deformation (or creep) to accommodate them [1-5]. Examples of such systems include powder compacts with non-uniform packing, adherent thin films on a rigid substrate and multilayered films. Furthermore, in the densification of powder compacts by pressure sintering (e.g., sinter-forging, hot pressing and hot isostatic pressing), the shrinkage is normally non-uniform so that shear stresses are developed [6]. Models for the consolidation process must take into account the volumetric strain as well as the shear strain [7-13]. It is therefore important not only to understand the densification and shear deformation (creep) processes separately but also their interaction.

Creep during the sintering of porous compacts has been investigated by De Jonghe, Rahaman and coworkers [14-18] and by Venkatachari and Raj and [19]. The approach consisted of applying a uniaxial load to a porous sintering powder compact and determine the volumetric strain rate and the creep rate separately from measurements of the axial shrinkage and radial shrinkage. For a given system, the ratio of the densification rate to the creep rate was found to be approximately constant over almost the entire sintering process (i.e., in the density range of ≈ 0.55 to 0.95 of the theoretical density). The magnitude of this ratio was also found to be dependent on the initial density of the powder compact and the initial particle size of the powder [17,18].

Densification and creep occur by a process of matter transport driven by chemical potential differences in the microstructure [20]. In densification, matter is removed from the grain boundaries (under compression due to an imposed hydrostatic stress) and deposited into the pores. Creep of dense solids typically occurs by matter transport from the boundaries under compression to those under tension. However, when matter transport occurs by diffusion, the shape change in creep must be accommodated by sliding of the grain boundaries [21]. Diffusion and grain boundary sliding occur sequentially so that the slower mechanism controls the rate of the process. Commonly in ceramic materials, sliding is believed to be the faster process so that diffusion is rate controlling. For a given system in which the material and experimental parameters (e.g., temperature, grain size and stress) are not significantly different, it is highly likely that the rate

controlling mechanism in densification and creep would be identical. In this case, changes in the densification behavior produced by changes in the diffusion coefficient would be expected to be reflected in the creep behavior.

The present work forms an attempt to provide further insight into densification, creep and their interaction. The approach is to modify the densification behavior of a given powder system (CeO_2) through the use of an additive (Y^{3+}) incorporated into solid solution and to determine the magnitude of the changes in the creep of dense specimens with similar compositions. CeO_2 has a simple cubic fluorite structure and does not undergo any known crystallographic phase changes during heating [22]. It also has a high solubility for many cations [23]. Additive cations have been shown to produce significant changes in the sintering behavior [24,25] and in the grain boundary mobility [26]. However, a systematic investigation of the effect of solid solution additives on the densification rate and the creep rate under similar conditions of temperature and atmosphere has not been performed.

2. EXPERIMENTAL PROCEDURE

2.1 Materials

Fine powders of undoped CeO_2 and CeO_2 solid solutions containing 2 at% Y and 6 at% Y were used in this study. They were prepared by precipitation from solution under hydrothermal conditions ($\approx 300^\circ\text{C}$ and $\approx 10\text{ MPa}$). The synthesis and characterization of the powders are described in detail elsewhere [24,25]. Briefly, X-ray diffraction (XRD) and transmission electron microscopy (TEM) revealed that the particles were crystalline, with an average size in the range of 10-15 nm. The powders were also found to be single phase with a cubic fluorite structure. No free Y was detected by XRD and, for a given powder, analytical TEM revealed no measurable difference in the concentration of Y from particle to particle.

The powders were dried for 5 h at 300°C and then compacted by uniaxial pressing in a stainless steel die (pressure $\approx 35\text{ MPa}$) followed by cold isostatic pressing ($\approx 400\text{ MPa}$). The density of the powder compacts was 0.51 ± 0.01 of the theoretical density.

2.2 Sintering

Compacts ($\approx 6\text{ mm}$ in diameter by $\approx 5\text{ mm}$) were sintered in O_2 in a dilatometer (1700 C; Theta Industries Inc., Port Washington, NY) that provided continuous monitoring of the shrinkage kinetics. Initially, sintering was performed at a constant heating rate of $5^\circ\text{C}/\text{min}$ to determine the temperature range of sinterability. To investigate the effect of Y on the densification rate, the compacts were sintered isothermally at 1150°C . In the experiments, the compacts were heated at $10^\circ\text{C}/\text{min}$ to the sintering temperature, held for 2 h and finally cooled at $30^\circ\text{C}/\text{min}$. The density

of the compact during sintering, ρ , was determined from the initial density, ρ_0 , and the measured shrinkage using the equation:

$$\rho = \rho_0 \left(1 - \frac{\Delta L}{L_0}\right)^3 \quad (1)$$

where L_0 is the initial sample length and $\Delta L = L_0 - L$, where L is the instantaneous sample length. In separate experiments, compacts were sintered for 0, 1 and 2 h at 1150 °C and used in the determination of the grain size as described later.

2.3 Creep Tests

To obtain dense specimens for creep testing, compacts (≈ 12 mm in diameter by ≈ 10 mm) were sintered in O_2 at 5 °C/min to 1400 °C and held at this temperature for 1 h. The final dimensions of the sintered specimens were ≈ 10 mm in diameter by ≈ 8 mm in length. The densities of the sintered specimens, measured from the mass and dimensions, were greater than 95% of the theoretical density. Additional compacts were sintered under the same conditions and used later in the determination of the grain size.

Compression creep tests were performed at a constant strain rate of 10^{-5} s^{-1} and 10^{-4} s^{-1} in air at 1200 °C in a modified MTS Model 810 mechanical tester. The change in length of the specimen was measured with a linear voltage displacement transducer (LVDT) and an associated electronic system capable of controlling the displacement rate to less than $10^{-6} \text{ }\mu\text{m/s}$. In the tests, the specimen was heated at 50 °C/min to the set point and held for 10 min to allow the temperature to stabilize. The load was then applied to produce the imposed strain rate. Both the load and strain were recorded automatically.

2.4 Characterization

The microstructure of the sintered specimens was observed from fractured surfaces in the scanning electron microscope (Leica Cambridge; Model 360FE). The average grain size was estimated from micrographs of the fractured surfaces by the linear intercept technique [27,28]. Approximately 100 grains were used in a given determination of the average grain size.

3. RESULTS

3.1 Densification

Figure 1 shows the data for the relative density, ρ , as a function of temperature, T , for the undoped CeO_2 (denoted CeO_2) and the CeO_2 solid solutions containing 2 at% Y ($\text{CeO}_2/2\text{Y}$) and 6

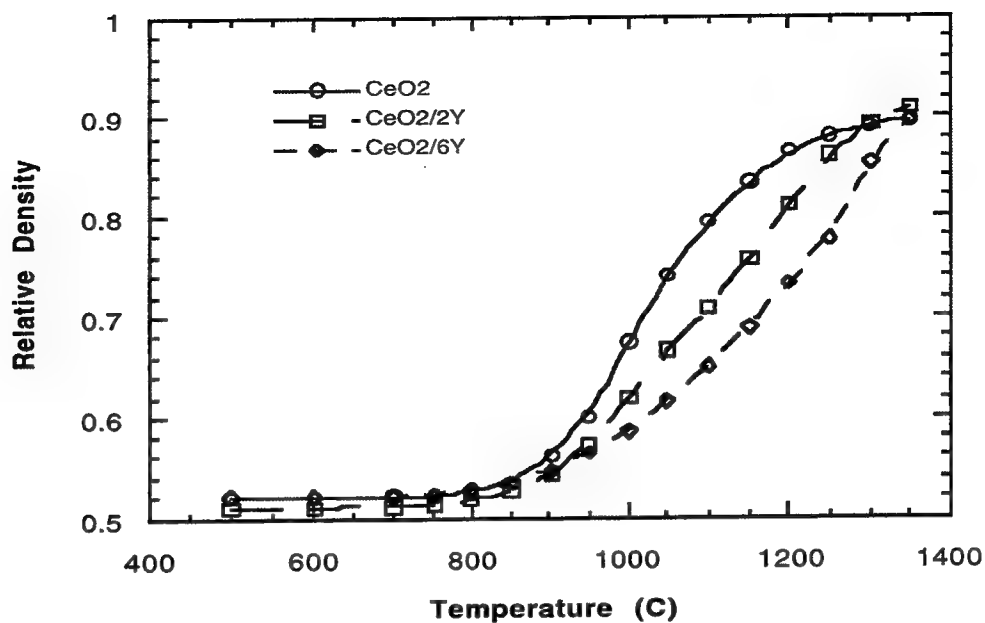


Figure 1. Relative density versus temperature for powder compacts of the undoped CeO₂ (denoted CeO₂) and the CeO₂ solid solutions containing 2 at% Y (CeO₂/2Y) and 6 at% Y (CeO₂/6Y) during constant heating rate sintering at 5 °C/min to 1350 °C in O₂.

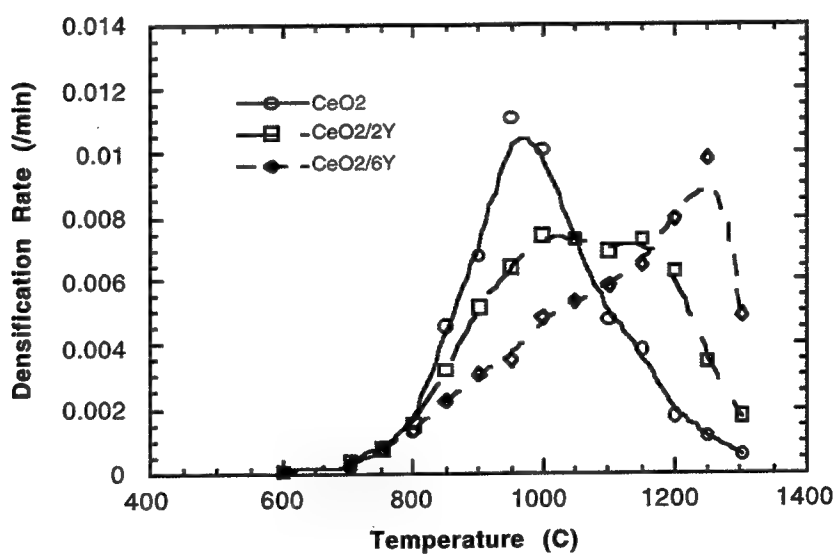


Figure 2. Densification rate as a function of temperature determined from the data in Fig. 1.

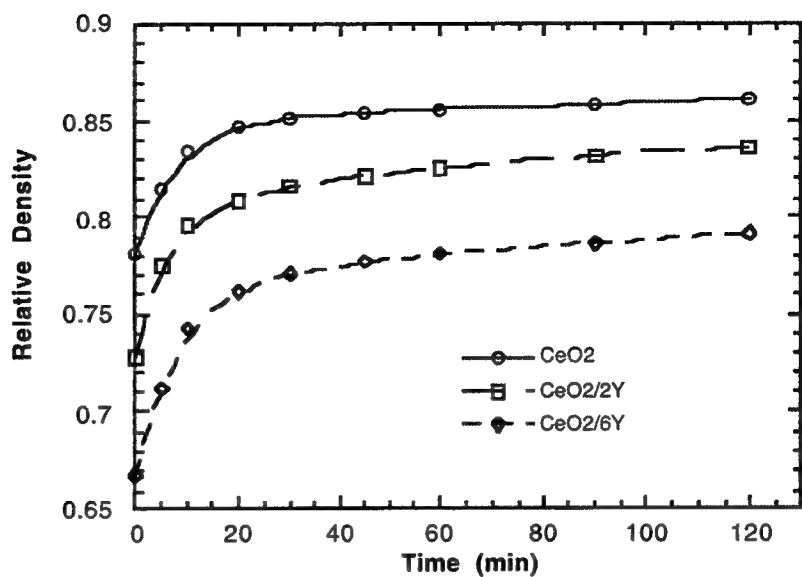


Figure 3. Relative density versus time for the CeO₂, CeO₂/2Y and CeO₂/6Y powder compacts during isothermal sintering at 1150 °C in O₂.

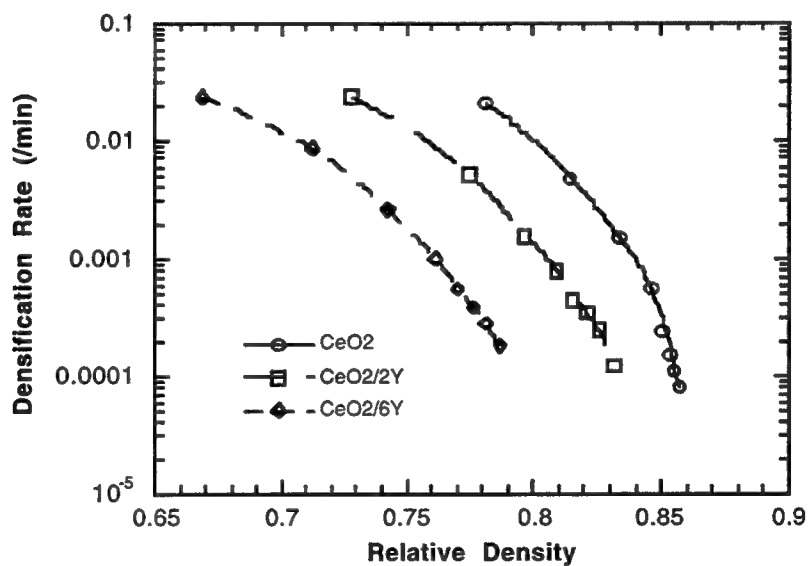


Figure 4. Densification rate as a function of relative density determined from the data in Fig. 3.

at% Y ($\text{CeO}_2/6\text{Y}$) during constant heating rate sintering at $5^\circ\text{C}/\text{min}$ to 1350°C . The compacts start to show measurable densification at $\approx 800^\circ\text{C}$. However, the influence of the Y additive is seen in its ability to suppress the densification process. By fitting smooth curves to the densification data and differentiating, the densification rate, $(1/\rho)d\rho/dt$, where t is the time, was determined as a function of T (Fig. 2). The maximum in the densification rate curve is shifted to higher temperatures with increasing concentration of Y. In constant heating rate sintering, both the density and the temperature (or diffusion coefficients) change during the experiment. To determine the effect of Y on the densification rate of CeO_2 , it is more convenient to utilize isothermal sintering data. The data of Figs. 1 and 2 show that an isothermal sintering temperature of $\approx 1150^\circ\text{C}$ would be expected to provide an adequate densification rate for meaningful comparison between the three specimens.

The data for ρ as a function of t during isothermal sintering at 1150°C are shown in Fig. 3. The curves have approximately the same shape and, at any time, the density is reduced with increase in the concentration of Y. The densification rate, found by fitting smooth curves to the data and differentiating, is shown in Fig. 4 as a function of ρ . The nearly similar shapes of the curves is an indication that the densification mechanism is not drastically affected by the presence of Y in solid solution. However, at any value of ρ , the densification rate is greatly reduced with increasing concentration of Y. Two main factors which can lead to this reduction in the densification rate are (i) the reduction in the diffusion coefficient for densification and (ii) an increase in the diffusion distance (i.e., grain size). It is more convenient to first investigate the effect of grain size.

Figure 5 shows scanning electron micrographs of the fractured surfaces of CeO_2 compacts after sintering for 0, 1 and 2 h at 1150°C . Because of the porous, fine-grained microstructure, it is difficult to make accurate measurements of the average grain size. However, the micrographs show clear evidence for appreciable grain growth, particularly after ≈ 1 h when the relative density increases very slowly above a value of ≈ 0.85 . This rapid increase in the grain size may be the cause of the noticeable change in the densification rate curve at $\rho \approx 0.85$. Corresponding micrographs for the CeO_2 solid solution containing 6 at% Y are shown in Fig. 6. (The increase in grain size with time for the $\text{CeO}_2/2\text{Y}$ specimen is somewhat intermediate between those for the CeO_2 and the $\text{CeO}_2/6\text{Y}$ specimens and the micrographs are omitted to reduce the length of the text.) The increase in grain size with time is much slower than that described above for the undoped CeO_2 . However, at a given time, there is a difference in density between the undoped CeO_2 and the CeO_2 containing 6 at% Y. A comparison of the grain size must be made at equivalent density.

Figure 7 shows a comparison of the scanning electron micrographs for (a) the undoped

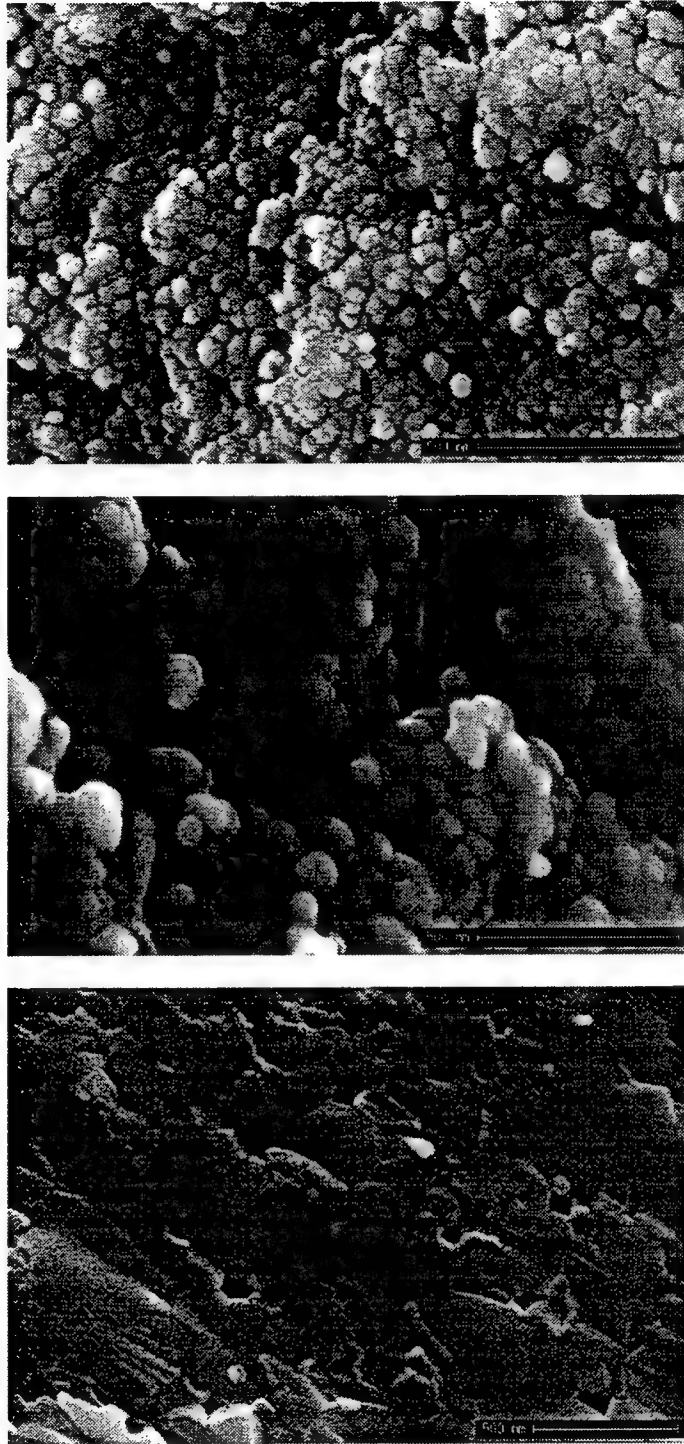


Figure 5. SEM of the fractured surfaces of the CeO_2 compact after sintering at 1150 °C for 0 h (top), 1 h (middle), and 2 h (bottom).

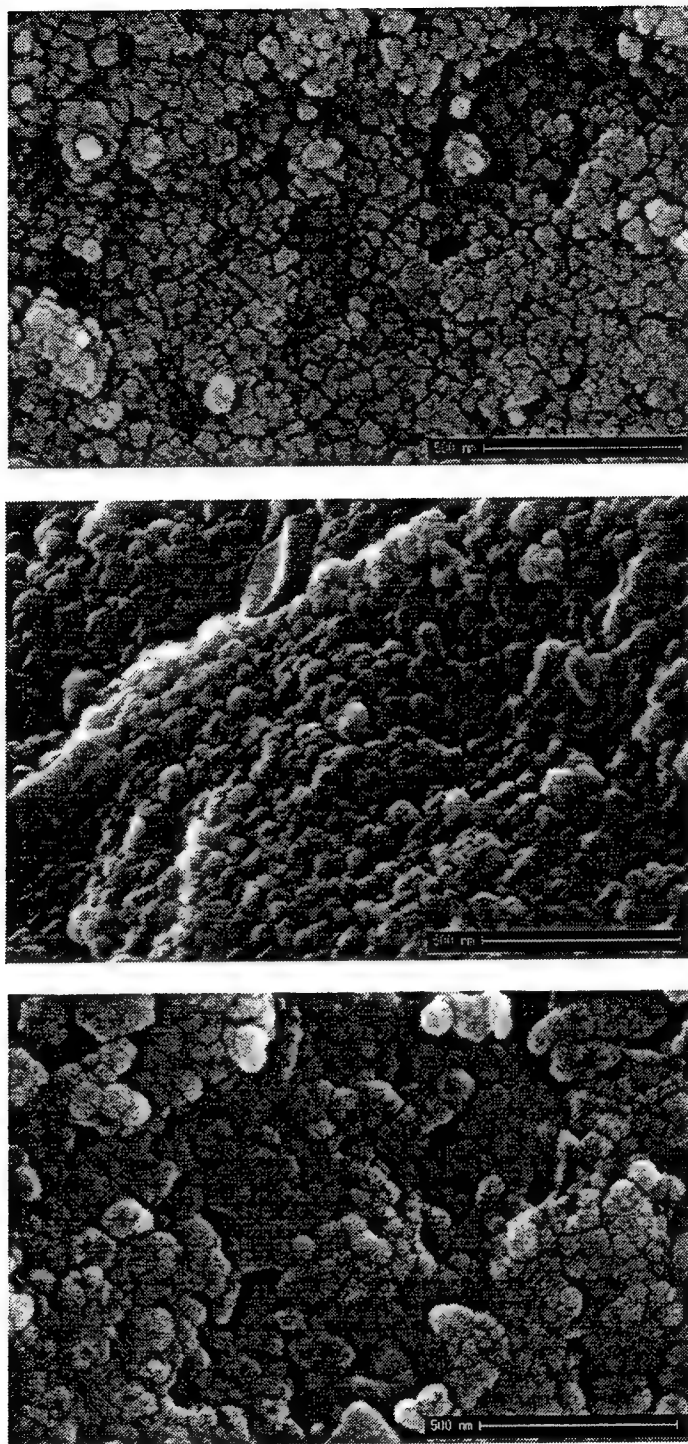


Figure 6. SEM of the fractured surfaces of the CeO₂/6Y compact after sintering at 1150 °C for 0 h (top), 1 h (middle), and 2 h (bottom).

CeO₂ with a relative density of 0.782, shown earlier as Fig. 5(a) and (b) the CeO₂/6Y specimen with a density of 0.786, shown earlier as Fig. 6(c). It is clear that at approximately the same density, the difference in grain size is negligible. It is therefore clear that the reduction in the densification rate produced by the presence of Y in solid solution is not due to a grain size effect.

3.2 Creep

Scanning electron micrographs of the fractured surfaces of the specimens (CeO₂, CeO₂/2Y and CeO₂/6Y) prior to creep testing are shown in Fig. 8. As outlined earlier, the specimens used in the creep tests were sintered under the same conditions (1 h at 1400 °C). The densities of the specimens, determined from the mass and dimensions, were approximately the same ($\rho = 0.95$ -0.97). The fracture mode is mainly transgranular. However, it is clear that the average grain size decreases significantly with increasing concentration of Y. In additional experiments, compacts were fractured and then sintered for 1 h at 1400 °C. Observations of the surface of the fracture (Fig. 9) revealed distinct grain boundaries. Micrographs similar to those shown in Fig. 9 were used to estimate the average grain size. The density and average grain size of the specimens used in the creep testing are summarized in Table 1.

For constant strain rates of 10^{-5} s^{-1} and 10^{-4} s^{-1} , the smoothed-average data for the applied stress as a function of the strain during creep at 1200 °C are shown in Fig. 10. At a given strain, the error in the stress is less than $\pm 10\%$ for the lower strain rate and less than $\pm 2\%$ for the higher strain rate. For the strain rate of 10^{-5} s^{-1} , a steady-state regime is reached above a strain of ≈ 0.01 where the stress required to sustain the imposed strain rate is approximately constant. A steady-state regime is reached at a higher stress for the higher strain rate. The data, particularly for the CeO₂ and CeO₂/6Y specimens, cover a narrow band. The steady-state stress is almost independent of the concentration of Y. However, as discussed in the next section, differences in grain size must be compensated for in order to determine the effect of the Y additive on the creep rate.

4. DISCUSSION

At an equivalent density and grain size, the isothermal sintering data (Fig. 5) show a dramatic effect of Y on the densification rate. For oxide ceramics, diffusion is expected to be the dominant mass transport mechanism during sintering [29,30]. Theoretical models for densification by diffusional mass transport predict an equation of the form [20]:

$$\dot{\epsilon}_p = \frac{A D \phi^{(m+1)/2}}{G^m kT} \Sigma \quad (2)$$

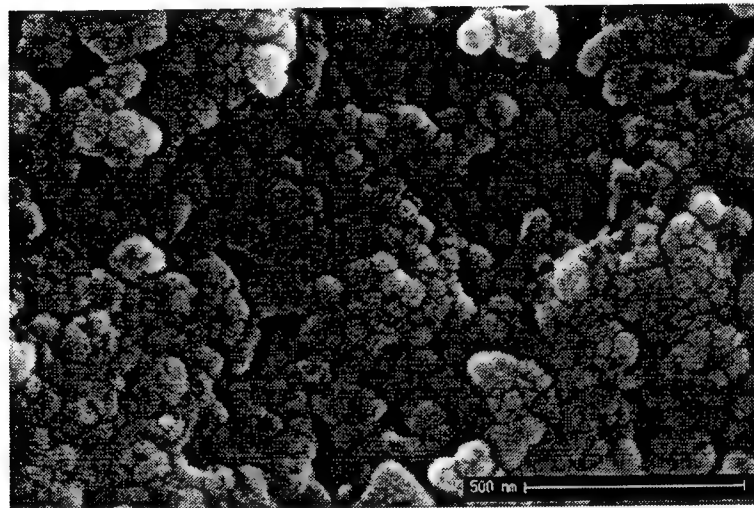
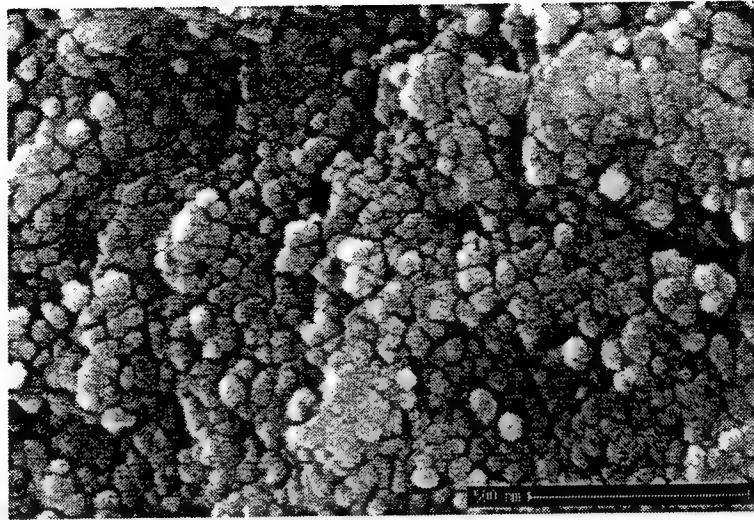


Figure 7. SEM of the fractured surfaces of the CeO₂ compact with a relative density of 0.782 (top) and the CeO₂/6Y compact with a relative density of 0.786 (bottom), showing approximately the same grain size at an equivalent density.

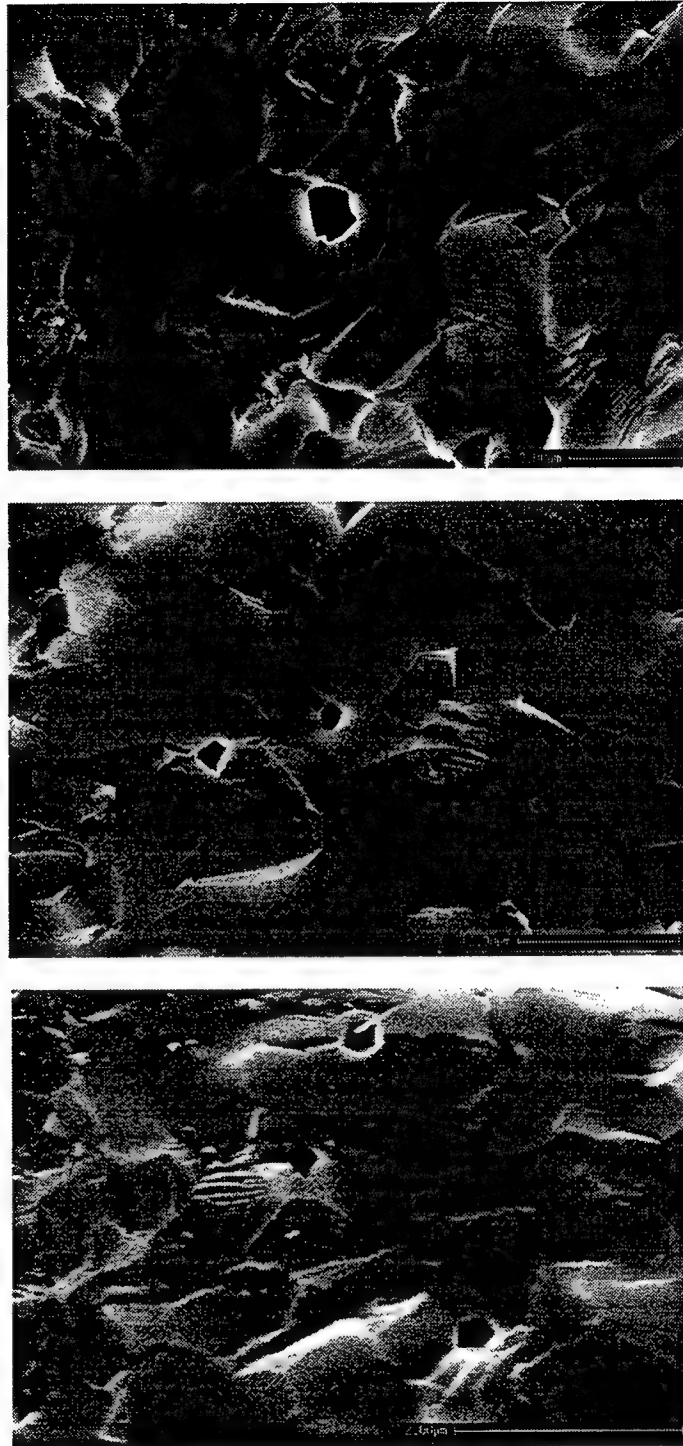


Figure 8. SEM of the fractured surfaces of the specimens used in creep testing: CeO₂ (top), CeO₂/2Y (middle), and CeO₂/6Y (bottom). The density and average grain size of the specimens are given in Table 1.

where $\dot{\epsilon}_D$ is the densification rate, A is a numerical constant, D is the diffusion coefficient for the rate-controlling densification mechanism, ϕ is the stress intensification factor, defined as the total cross-sectional area divided by the effective grain boundary area, Σ is the sintering stress, defined as the equivalent external stress which would have the same effect on the densification as the pores and grain boundaries, G is the grain size, k is the Boltzmann constant, T is the absolute temperature and m is an exponent that depends on the mechanism of densification ($m = 3$ for grain boundary diffusion and $m = 2$ for lattice diffusion). At the equivalent density and grain size, the parameters ϕ and Σ are not expected to be affected strongly by the presence of the Y in solid solution. Therefore, the drastic reduction in the densification rate produced by the Y additive appears to be due to a decrease in the diffusion coefficient. While the mechanism of densification of CeO_2 is not clear, predictions based on Herring's scaling laws [31] indicate that grain boundary diffusion is enhanced with decreasing particle size. Compared to lattice diffusion, grain boundary diffusion is also expected to be more rapid at lower temperatures. For the fine powders and relatively low sintering temperatures used in this work, it appears reasonable to assume that grain boundary diffusion is the rate-controlling mechanism.

For the creep of porous solids by diffusion, theoretical models predict an equation of the form [20]:

$$\dot{\epsilon}_C = \frac{B D \phi^{(m+1)/2} p_z}{G^m kT} \quad (3)$$

where $\dot{\epsilon}_C$ is the creep rate, B is a numerical constant, D is the diffusion coefficient for the rate-controlling creep mechanism, and p_z is an applied uniaxial stress. The specimens used in the creep tests had a fairly high density (Table 1) so that $\phi \approx 1$. For the creep of dense specimens at a given temperature, Eq. (3) predicts that the stress required to sustain a fixed strain rate varies as G^m/D . As the data in Table 1 also show, the grain size of the specimens decrease significantly with increasing concentration of Y. However, the data (Fig. 10) show little variation in the steady-state creep stress. It follows, therefore, that the Y additive must act to reduce the diffusion coefficient.

The determination of the creep mechanism and, hence, the value of m in Eq. (3) requires further work. However, as shown in Fig. 11, assuming a mechanism of grain boundary diffusion ($m = 3$) leads to a reduction in D which is comparable to that observed in densification. For a mechanism of lattice diffusion ($m = 2$), the reduction in D is still significant but somewhat smaller than that observed in densification. From a consideration of the low stresses and strain rates used in the creep tests, the relatively fine grain size of the specimens, and a test temperature close to that used in the sintering experiments, a strong argument may be made for the creep and densification

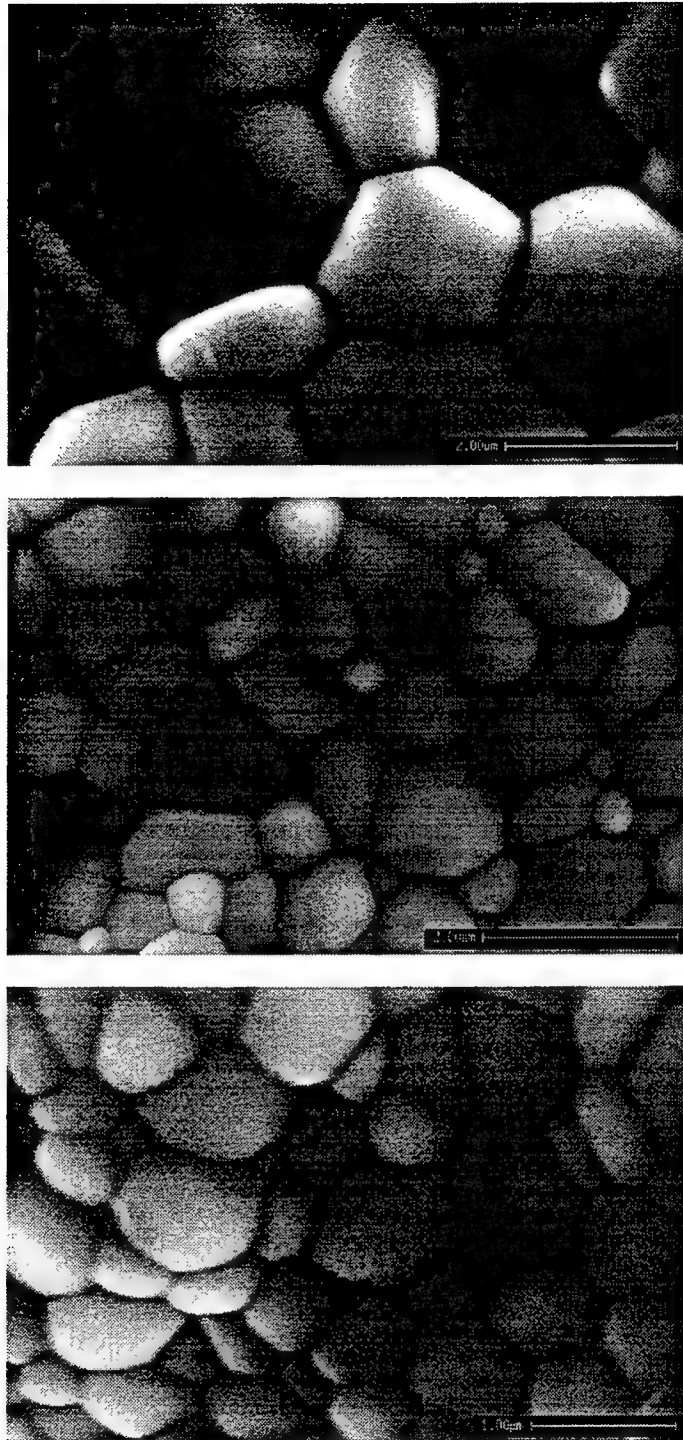


Figure 9. SEM of the surfaces of cracked specimens used in the determination of the average grain size of the creep specimens.

Table 1. Relative density and average grain size of the specimens used in the creep tests.

Specimen	Relative Density	Grain Size (μm)
CeO ₂	0.95	3.5
CeO ₂ /2 at% Y	0.97	2.0
CeO ₂ /6 at% Y	0.96	0.75

mechanisms being identical. As outlined earlier, this mechanism appears to be grain boundary diffusion.

The comparable effects on the densification and creep processes produced by the Y additive in solid solution strongly indicate that it may be possible to predict the long-term creep behavior from short-term sintering experiments. The ability to predict the creep behavior from the densification behavior can provide a distinct advantage in that creep experiments are fairly time-consuming and require more sophisticated equipment. The results also indicate that the solid solution approach, in which cations are incorporated into solid solution to provide structural modification, may be a valid approach for controlling the creep behavior of monolithic ceramics. This approach, commonly used in metals, has received little attention in ceramic materials for structural applications.

5. CONCLUSIONS

The use of Y³⁺ as a solid solution additive produces a drastic reduction in the densification rate of CeO₂ powder compacts. At a concentration of 6 at% Y, the densification rate at a given density and grain size is reduced by a factor of ≈ 100 during isothermal sintering at 1150 °C. This reduction can be explained in terms of a reduction in the diffusion coefficient for the densification process. In the creep of dense specimens at a comparable temperature (1200 °C), Y³⁺ also produces a significant reduction in the grain size compensated creep rate. For a creep mechanism of grain boundary diffusion, which is expected to be the dominant mechanism in densification and creep under the present conditions, the reduction in the creep rate produced by the Y additive is comparable to the observed reduction in the densification rate. The closely similar effects in densification and creep produced by the Y additive indicate that it may be possible to predict long-term creep behavior from short-term sintering experiments provided that both processes are controlled by the same mechanism. The use of solid solution additives may provide a valid approach for controlling the creep resistance of monolithic ceramics, particularly rare earth oxides that have a high solid solubility for many cations.

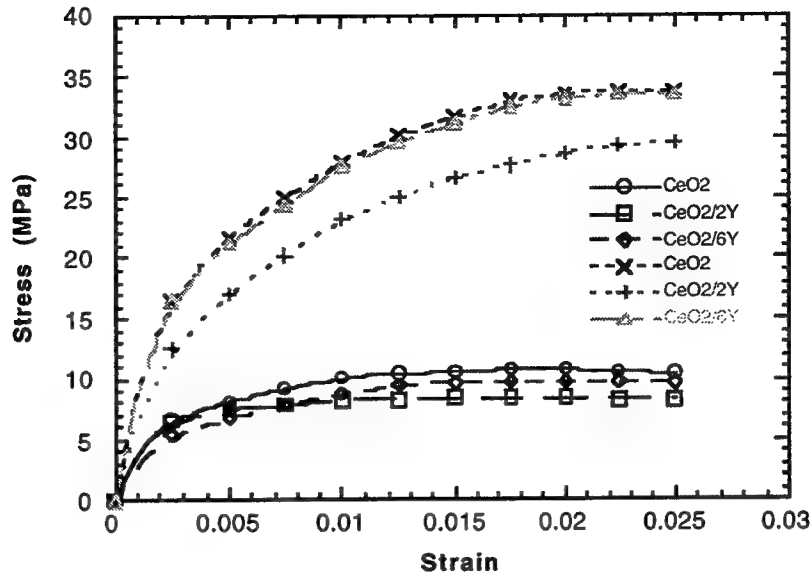


Figure 10. True stress as a function of true strain during creep testing at constant strain rates of 10^{-4} s^{-1} (upper curves) and 10^{-5} s^{-1} (lower curves) for the CeO₂, CeO₂/2Y and CeO₂/6Y specimens. The tests were performed in air at 1200 °C.

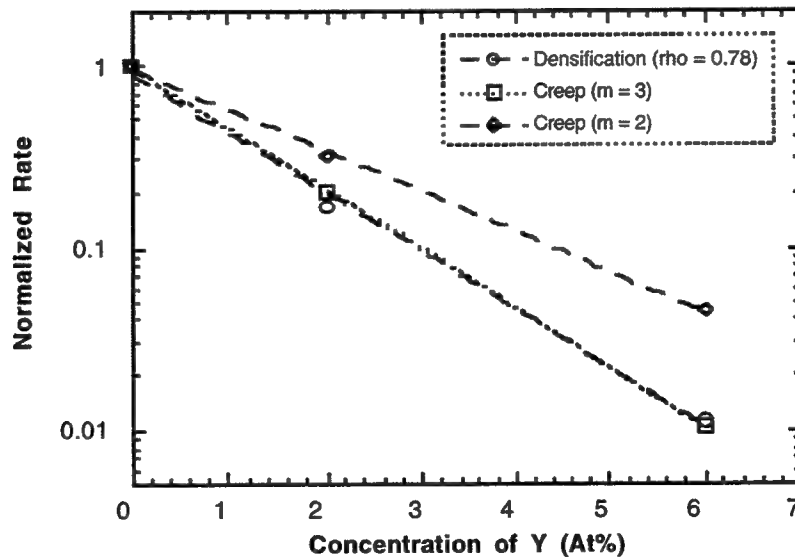


Figure 11. Grain size compensated densification rates (at a fixed relative density of 0.78) and creep rates, normalized to that for the undoped CeO₂, versus the concentration of Y. The creep rates were determined separately for assumed mechanisms of grain boundary diffusion ($m = 3$) and lattice diffusion ($m = 2$).

REFERENCES

1. A. G. Evans, *J. Am. Ceram. Soc.*, **65** 497 (1982).
2. L. C. De Jonghe, M. N. Rahaman, and C. H. Hsueh, *Acta metall.*, **34** 1467 (1986).
3. R. Raj and R. K. Bordia, *Acta metall.*, **32** 1003 (1984).
4. C. H. Hsueh, A. G. Evans, R. M. Cannon, and R. J. Brook, *Acta metall.*, **34** 927 (1986).
5. G. W. Scherer, *J. Am. Ceram. Soc.*, **70** 719 (1987).
6. W. B. Li, M. F. Ashby, and K. E. Easterling, *Acta metall.*, **35** 2831 (1987).
7. R. E. Dutton, S. Shamasundar, and S. L. Semiatin, *Metall. Mater. Trans.*, **26A** 2041 (1995).
8. J. Besson and M. Abouaf, *J. Am. Ceram. Soc.*, **75** 2165 (1992).
9. J. Besson, F. Valin, P. Lointier, and M. Boncoeur, *J. Mater. Sci. Perf.*, **1** 637 (1992).
10. Z.-Z. Du and A. C. F. Cocks, *Acta metall. mater.*, **40** 1969 (1992).
11. Z.-Z. Du and A. C. F. Cocks, *Acta metall. mater.*, **40** 1981 (1992).
12. D. M. Elzey and H. N. G. Wagley, *Acta metall. mater.*, **41** 2297 (1993).
13. Y.-M. Liu, H. N. G. Wadley, and J. M. Duva, *Acta metall. mater.*, **42** 2247 (1994).
14. M. N. Rahaman, L. C. De Jonghe, and R. J. Brook, *J. Am. Ceram. Soc.*, **69** 53 (1986).
15. M. N. Rahaman and L. C. De Jonghe, *J. Am. Ceram. Soc.*, **73** 707 (1990).
16. R. E. Dutton and M. N. Rahaman, *J. Am. Ceram. Soc.*, **75** 2146 (1992).
17. M.-Y. Chu, M. N. Rahaman, and L. C. De Jonghe, *J. Am. Ceram. Soc.*, **74** 514 (1991).
18. M.-Y. Chu, M. N. Rahaman, and L. C. De Jonghe, *Acta metall.*, **37** 1415 (1989).
19. K. R. Venkatachari and R. Raj, *J. Am. Ceram. Soc.*, **69** 499 (1986).
20. M. N. Rahaman, *Ceramic Processing and Sintering*, Marcel Dekker, New York (1995).
21. R. Raj and M. F. Ashby, *Acta metall.*, **23** 63 (1975).
22. R. N. Blumental, F. S. Brugner, and J. E. Garnier, *J. Electrochem. Soc.*, **120** 1230 (1973).
23. R. T. Distine, R. N. Blumental, and T. F. Kuech, *J. Electrochem. Soc.*, **126** 264 (1979).
24. Y.-C. Zhou and M. N. Rahaman, *J. Mater. Res.*, **8** 1680 (1993).
25. M. N. Rahaman and Y.-C. Zhou, *J. Europ. Ceram. Soc.*, **15** 939 (1995).
26. P.-L. Chen and I.-W. Chen, *J. Am. Ceram. Soc.*, **77** 2289 (1994).
27. M. I. Mendelson, *J. Am. Ceram. Soc.*, **52** 443 (1969).
28. J. C. Wurst and J. A. Nelson, *J. Am. Ceram. Soc.*, **55** 109 (1972).
29. M. F. Ashby, *Acta metall.*, **22** 275 (1974).
30. F. B. Swinkels and M. F. Ashby, *Acta metall.*, **29** 259 (1981).
31. C. Herring, *J. Appl. Phys.*, **21** 301 (1950).

AB INITIO MODELING OF THE ENTHALPIES
OF FORMATION OF FLUOROCARBONS

Martin Schwartz
Regents Professor
Department of Chemistry

University of North Texas
215 W. Sycamore
Denton, TX 76203

Final Report for:
Summer Faculty Research Program
Materials Directorate
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

October 1996

AB INITIO MODELING OF THE ENTHALPIES OF FORMATION OF FLUOROCARBONS

Martin Schwartz
Regents Professor
Department of Chemistry
University of North Texas

Abstract

The *ab initio* G2, G2(MP2), CBS-4 and CBS-Q quantum mechanical protocols and the parameterized BAC-MP4 procedure were used to calculate the enthalpies of formation ($\Delta_f H^\circ$) of ethane and the complete series of fluoroethanes, $C_2H_xF_{6-x}$, $x=0-5$. Results from all methods exhibited significant negative deviations from experiment. With the exception of the CBS-4 and BAC-MP4 procedures, the negative errors in the calculated enthalpies were observed to be linearly dependent upon the number of C-F bonds in the molecule. Application of a Bond Additivity Correction (BAC) parameter, Δ_{CF} , derived in an earlier investigation of fluoro- and chlorofluoromethanes, removed some, although not all of the systematic deviations. Introduction of a heavy atom interaction parameter, representing the effect of an attached carbon on the C-F bond error yielded corrected enthalpies which agree with experiment to within the reported uncertainties. The BAC-MP4 method, which has already been parameterized with generalized BAC's, yields calculated enthalpies which average approximately 10 kJ/mol below the experimental values of $\Delta_f H^\circ$ in the fluoroethanes.

The G2 and G2(MP2) quantum mechanical procedures have been used to calculate the enthalpies of formation ($\Delta_f H^\circ$) of ethylene, acetylene and all of their fluorinated derivatives, $C_2H_xF_{4-x}$, $x=0-4$ and $C_2H_xF_{2-x}$, $x=0-2$. Values obtained from both methods exhibit large, systematically negative, deviations from experiment, with errors that increase approximately linearly with the number of C-F bonds. For each method, application of a C-F bond additivity correction, Δ_{CF} , and heavy atom interaction parameter, Δ_{HH} , whose values were derived from earlier studies of saturated HFC's, reduced the errors dramatically; the resulting RMS deviations from experiment are lower than the RMS uncertainties in the measured enthalpies. The only reported enthalpy of formation for fluoroacetylene, obtained from average bond enthalpies, has an extremely large uncertainty (> 60 kJ/mol). Based upon the overall good agreement of the calculated heats of formation with experiment, we recommend $\Delta_f H^\circ(H-C\equiv C-F) = 114 \pm 10$ kJ/mol.

AB INITIO MODELING OF THE ENTHALPIES OF FORMATION OF FLUOROCARBONS

Martin Schwartz

TABLE OF CONTENTS

	page
CHAPTER 1. AB INITIO CALCULATIONS OF THE ENTHALPIES OF FORMATION OF FLUOROETHANES	4
A. Introduction	4
B. Computational Methods	4
C. Results and Discussion	4
D. Bond Additivity Corrections	5
E. Summary and Conclusions	7
F. References	8
 CHAPTER 2. CALCULATED ENTHALPIES OF FORMATION OF FLUOROETHYLENES AND ACETYLENES	 10
A. Introduction	10
B. Calculations	10
C. Results and Discussion	11
D. References	12
 TABLES AND FIGURES	
Table 1. Enthalpies of Formation in Fluoroethanes	14
Table 2. Bond Additivity Correction Parameters in Fluoroethanes	16
Table 3. Enthalpies of Formation in Fluoroethylenes and Acetylenes	17
 Figure 1. Errors in Enthalpies of Formation of Fluoroethanes	 18
Figure 2. Errors in Enthalpies of Formation of Fluoroethylenes and Acetylenes	19

CHAPTER 1. INITIO CALCULATIONS OF THE ENTHALPIES OF FORMATION OF FLUOROMETHANES.

INTRODUCTION

Because of their lower ozone depletion potential, C_1 and C_2 hydrofluorocarbons (HFC's) have been proposed as interim replacements for the chlorofluorocarbons (CFC's) and halon fire suppressants [1]. In order to enable accurate modeling of the fire suppression capabilities of HFC's and other potential halon replacements, we have recently performed *ab initio* quantum mechanical calculations of the enthalpies of formation ($\Delta_f H^\circ$) of fluoromethanes [2] and chlorofluoromethanes [3,4].

In this study, we extend these calculations to the C_2 HFC's and report the results of G2(MP2), G2, CBS-4, CBS-Q and BAC-MP4 calculations of enthalpies of formation of ethane and the complete series of fluoroethanes, $C_2H_xF_{6-x}$, $x=0-5$. The results are presented below.

COMPUTATIONAL METHODS

The *ab initio* calculations were performed using the GAUSSIAN code [5-7] on HP-PARisk, CRAY Y-MP, CRAY X-MP, SUN-Sparc and SGI Power-Challenge computers [8]. Details of these calculations are described in an earlier paper [2] as well as the original articles that pertain to the BAC-MP4 [9-11], G2 [12-14], and CBS family of methods [15-18].

RESULTS AND DISCUSSION

The reported experimental enthalpies of formation of the fluoroethanes with their estimated uncertainties were taken from refs. [21-24] and are presented in the second column of Table 1A. The remaining columns of the table contain the enthalpies calculated by the various *ab initio* methods. The quantities in parentheses represent the deviations from experiment. The deviations in the *ab initio* enthalpies are also given in category A of Table 2 and plotted in Fig. 1 for the G2 and G2(MP2) methods [closed circles and solid lines].

One observes from the RMS and average errors in the table that the G2(MP2), G2 and CBS-Q enthalpies exhibit large negative deviations from experiment. Furthermore, from Figs. 1A, 1B and 1D, one sees that these negative errors are systematic with an approximately linear dependence upon the number of C-F bonds in the molecule [25]. The CBS-4 enthalpies also show negative deviations from the experimental values. However, for this method the errors are

approximately constant.

The four *ab initio* procedures have been found to yield very accurate enthalpies of formation in most non-halogenated species. From these results, and from previous work on fluorinated [2] and chlorinated [3] hydrocarbons it is clear that these methods exhibit severe, unacceptable errors relative to reported experimental enthalpies.

BOND ADDITIVITY CORRECTIONS

When the errors in $\Delta_f H^\circ$ are observed to be systematically dependent upon the number of bonds of a given type in a molecule, they can in principle be removed by the application of Bond Additivity Corrections (BAC's), originally introduced by Melius in his BAC-MP4 method [9-11]. The correction, Δ_{BAC} , is of the form: $\Delta_{BAC} = \Delta_{C-X} n_{C-X}$, where n_{C-X} is the number of a given type of bond and Δ_{C-X} is the BAC parameter representing the error associated with this bond type. The parameters are determined via a linear fit of Eq. [1] to the calculated deviations of the *ab initio* enthalpies from experiment. In our earlier investigation on thermochemistry of the chlorofluoromethanes [3], it was found that, for the same four *ab initio* methods, the application of C-F and C-Cl BAC parameters, removed all systematic error and yielded corrected enthalpies of formation whose RMS deviations from experiment were significantly below the RMS uncertainties in the literature data. It was also observed that values of Δ_{C-H} were smaller than the parameter's standard deviation and that its exclusion had no impact on the quality of the fit.

Because of the close similarity of the fluoroethanes (C2's) to the fluoromethanes (C1's) in the earlier work, it would be reasonable to expect that the same C-F BAC parameter should also correct the systematic errors found in this study. The results of this procedure are displayed in category B of Table 2. The second column contains the values of Δ_{C-F} from Ref. [3], with their estimated standard deviations. The next two pairs of columns represent the RMS and average errors for the C1's and C2's, respectively. As noted above, the values of RMS1 and AVE1 demonstrate that the application of a single C-F BAC does indeed remove virtually all systematic errors in $\Delta_f H^\circ$ for the fluoromethanes; note that resulting RMS deviations are significantly smaller than the RMS experimental uncertainty of ± 4.1 kJ/mol. From RMS2 in Table 2, one sees, of course, that the same BAC also yields a substantial decrease in RMS deviations in the C2's. However, in contrast to the C1's, the comparatively large negative values of AVE2 reveals that not all of the negative systematic error has been removed by the bond additivity correction. Furthermore, the

resultant values of RMS2 are, in all cases, significantly greater than the RMS experimental uncertainty of ± 4.8 kJ/mol. In principle, this problem could be remedied by refitting to minimize the RMS residuals in the fluoroethanes. However, this yields a larger value for the parameter, which would overcorrect enthalpies in the C1's. Further, it is physically unrealistic that the error due to a C-F bond should differ in the two series.

An alternative approach to eliminate the remaining systematic error would be via the addition of a second BAC, Δ_{C-C} , to account for the presence of a C-C bond in the fluoroethanes. However, this has the effect of increasing the error in ethane, where the *ab initio* enthalpies are in excellent agreement with experiment for all but the CBS-4 method (Table 1A). Furthermore, these methods have been applied with great success to the prediction of numerous non-fluorinated organic species with one or more C-C bonds. Hence, the introduction of a C-C bond additivity correction would not be consistent with these earlier results.

In order to explore further the source of the remaining systematic errors in the fluoroethanes, we refer to the results of our earlier investigation of the chlorofluoromethanes [3]. In that work, it was discussed at some length how trends and curvature in plots of (calc) - (expt) vs. n_{C-F} (at fixed n_{C-Cl}) and vs. n_{C-Cl} (at fixed n_{C-F}) provided definitive evidence of "heavy atom" interactions (also utilized by Melius in the BAC-MP4 method [9-11]); i.e. the presence of one C-F bond increases the error due to either a second fluorine or a chlorine atom on the same carbon. In the earlier work, it was decided not to include heavy atom interactions since the RMS residuals using linearly independent BAC's were already below the experimental uncertainties in the chlorofluoromethanes. In contrast, for the fluoroethanes, the comparatively large residual errors (RMS2) and negative average deviations (AVE2) indicate that the introduction of a heavy atom interaction parameter is necessary in this series to account for enhanced errors due to the presence of a second carbon atom attached to the carbon containing the C-C bond. In the spirit of the BAC-MP4 procedure, one may write: $\Delta_{BAC} = n_{C-F}\Delta_{C-F}f_C$, where f_C is the interaction parameter representing the impact of a second attached carbon on the error due to the C-F bond. This has no impact on the C1's since $f_C = 1$ if there is no second carbon.

One may solve for f_C via linear regression (constrained to pass through the origin) to fit Δ_{BAC} to the deviations of the *ab initio* enthalpies from experiment, with the product $n_{C-F}\Delta_{C-F}$ as the independent variable. The resultant values of the interaction parameter are shown in the third column of category C in Table 2. The estimated errors in f_C may be obtained from either (a) the

standard deviation of the parameter obtained from the least squares fit, or (b) the variation in the parameter when σ is varied between its upper and lower limits ($\pm \sigma$). The estimates given in the table represent the larger of these two values. The very large value of $\sigma()$ for the CBS-4 method is a consequence of the large error in σ for this method and the fact that, as mentioned above, errors in the CBS-4 enthalpies of formation are largely random.

It is very satisfying to observe from RMS2 in category C that the RMS residuals using a C-F BAC with an interaction parameters are reduced significantly for all but the CBS-4 method. Further, the much smaller negative values of AVE2, relative to category B, indicates that almost all of the remaining systematic error has been removed from the BAC corrected enthalpies of formation. With the exception of the CBS-4 method, values of RMS2 have been reduced to values comparable to the RMS experimental uncertainty. The removal of systematic error is also demonstrated in Fig. 1, in which, with the exception of the CBS-4 method, errors in the BAC corrected enthalpies of formation (open circles and dashed line) are clustered about zero.

Individual BAC corrected enthalpies of formation in the fluoromethanes, using from Ref. [3] and from this work (category C of Table 2) are presented in Table 1B. It is of interest to note that BAC-G2, BAC-G2(MP2) and BAC-CBS-Q enthalpies agree with experiment to within 10 kJ/mol (and are usually much closer), except for $\text{CH}_2\text{FCH}_2\text{F}$, where all three methods exhibit markedly negative deviations from the reported value. Since the reported uncertainty in the experimental enthalpy of this compound is quite large (Table 1A), it is tempting to speculate that the literature value may be in error. If this compound is excluded from the statistics, then all three methods show RMS errors significantly under 5 kJ/mol, and average errors under 1 kJ/mol.

Finally, the last column of Table 1B contains errors in enthalpies of formation of the fluoroethanes calculated using the Melius BAC-MP4 procedure [9-11]. It was shown in Ref. [3] that calculated enthalpies using this method are in excellent agreement with experimental results for the fluoromethanes. There are, however, rather large residual systematic errors in predicted values of $\Delta_f H^\circ$ of the fluoroethanes, as illustrated by the RMS and average errors of 11.5 and -10.2 kJ/mol, respectively. From the table, it is seen that calculated enthalpies of the C2 series are uniformly more negative than results from the literature. Thus, it would appear that the C-C BAC and interaction parameters derived in the BAC-MP4 method to fit data in predominantly non-halogenated organics do not adequately characterize the systematic errors in fluoroethanes.

SUMMARY AND CONCLUSIONS

Enthalpies of formation of the fluoroethanes calculated by the G2, G2(MP2), CBS-4 and CBS-Q quantum mechanical procedures were found to exhibit substantial negative deviations from experiment. With the exception of the CBS-4 technique, where the deviations are random, the errors are proportional to the number of C-F bonds in the molecule. The application of a C-F bond additivity correction, , derived in an earlier investigation of fluoro- and chlorofluoromethanes, removed some, although not all of the systematic errors. Introduction of a further heavy atom interaction parameter, , to account for the impact of a second attached carbon on the C-F bond error, yielded BAC corrected enthalpies with RMS errors comparable to experimental uncertainties.

Additional calculations on fluoroethylenes and acetylenes, fluorocarbon radicals and reactive transition states are in progress to assess the capability of *ab initio* procedures with bond additivity corrections to predict accurate enthalpies of formation in these species.

REFERENCES

- [1] Banks, R. E. *J. Fluorine Chem.* **1994**, *67*, 193
- [2] Berry, R. J.; Burgess, D. R., Jr.; Nyden, M. R.; Zachariah, M. R.; Schwartz, M. "Halon Thermochemistry: *Ab Initio* Calculations of the Enthalpies of Formation of Fluoromethanes," *J. Phys. Chem.* **1995**, *99*, 17145.
- [3] Berry, R. J.; Burgess, D. R., Jr.; Nyden, M. R.; Zachariah, M. R.; Schwartz, M. "Halon Thermochemistry: Calculated Enthalpies of Formation of Chlorofluoromethanes," *J. Phys. Chem.* **1996**, *100*, 7405.
- [4] The systems investigated in Ref. [3] include methane, the four fluoromethanes, the four chloromethanes, and the six mixed species.
- [5] Gaussian 90, Revision F, Gaussian, Inc.: Pittsburgh, PA, **1990**.
- [6] *Gaussian 92*, Revision F.4., Gaussian, Inc.: Pittsburgh, PA, **1992**.
- [7] *Gaussian 94*, Revision A.1, *Gaussian, Inc.*, Pittsburgh, PA, **1995**.
- [8] Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. In no case does such identification imply recommendation or endorsement.
- [9] Melius, C. F. Thermochemistry of Hydrocarbon Intermediates in Combustion. Applications of the BAC-MP4 Method. In *Springer-Verlag DFVLR Lecture Notes*;

Springer-Verlag: Berlin, 1990.

- [10] Melius, C. F. Thermochemical Modeling I. Application to Ignition and Combustion of Energetic Materials. *Chemistry and Physics of Energetic Materials*; S. N. Kluwer Academic: New York, 1992.
- [11] Ho, P.; Melius, C. F. *J. Phys. Chem.* **1990**, *94*, 5120
- [12] Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622.
- [13] Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A.; *J. Chem. Phys.* **1990**, *93*, 2537.
- [14] Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; and Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.
- [15] Petersson, G. A.; Al-Laham, M. A. *J. Chem. Phys.* **1991**, *94*, 6081
- [16] Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* submitted.
- [17] Nyden, M. R.; Petersson, G. A. *J. Chem. Phys.* **1981**, *75*, 1843. (b) Petersson, G. A.; Nyden, M. R. *J. Chem. Phys.* **1981**, *75*, 3423. (c) Petersson, G. A.; Licht, S. L. *J. Chem. Phys.* **1981**, *75*, 4556. (d) Petersson, G. A.; Yee, A. K.; Bennett, A. *J. Chem. Phys.* **1985**, *83*, 5105. (e) Petersson, G. A.; Braunstein, M. *J. Chem. Phys.* **1985**, *83*, 5129.
- [18] Petersson, G. A.; *Proc. Nat. Acad. Sci. (USA)* **1974**, *71*, 2795
- [19] Durig, J. R.; Liu, J.; Little, T. S.; Kalasinsky, V. F. *J. Phys. Chem.* **1992**, *96*, 8224.
- [20] Kalasinsky, V. F.; Anjaria, H. V.; Little, T.S. *J. Phys. Chem.* **1982**, *86*, 1351.
- [21] Tsang, W.; Hampson, R. F. *J. Phys. Chem. Ref. Data* **1986**, *15*, 1087.
- [22] Chen, S. S.; Rodgers, A. S.; Chao, J.; Wilhoit, R. C.; Zwolinski, B. J. *J. Phys. Chem. Ref. Data* **1975**, *4*, 441.
- [23] Burgess, D. R., Jr.; Zachariah, M. R.; Tsang, W.; Westmoreland, P. R. 'Thermochemical and Chemical Kinetic Data for Fluorinated Hydrocarbons.' *Prog. Energy Combust. Sci.* **1996**, *21*, 453.
- [24] Lacher, J. A.; Skinner, H. A. *J. Chem. Soc. (A)* **1968**, 1034.
- [25] Fits of $\Delta_f H^\circ(\text{calc}) - \Delta_f H^\circ(\text{expt})$ vs. n_{CF} to second order polynomials yielded statistically

insignificant quadratic terms.

- [26] The standard errors in Δ_{CH} and Δ_{CF} (Table 2) are equal because of the multicollinearity of the two variables, n_{CH} and n_{CF}
- [27] One observes from category B of Table 6 that, for all four methods, Δ_{CH} is relatively small and comparable to or less than the standard error in this parameter.
- [28] Kolesov, V. P. *Russ. Chem. Rev.* **1978**, *47*, 599.
- [29] For all four *ab initio* methods, the slopes of the plots of $\Delta_f H^0(\text{BAC}) - \Delta_f H^0(\text{expt})$ vs. n_{CF} [dashed lines in Fig. 1] are smaller than their standard deviations.

CHAPTER 2. Calculated Enthalpies of Formation of Fluoroethylenes and Acetylenes

INTRODUCTION

Because of their stratospheric ozone depletion potential [1], manufacture of currently used halon fire suppressants (e.g. CF_3Br , CF_2ClBr , $\text{CF}_2\text{BrCF}_2\text{Br}$) has been banned under the Montreal Protocols [2]. In order to obtain the requisite data to model the mechanism of flame suppression by potential halon replacements, we have utilized *ab initio* quantum mechanics to predict accurate thermochemical data [3-5] and rate constants [6] for the reactions of HFC's and CFC's and their decomposition products.

We have recently completed investigations on the application of various current methods [G2, G2(MP2), CBS-Q and CBS-4] to calculate the enthalpies of formation [$\Delta_f H^0$] of fluoro-, chloro- and chlorofluoromethanes and fluoroethanes [3-5]. Fluoroethylenes are important intermediates in the flame chemistry of HFC's [7,8]. Hence, we have chosen to extend the G2 and G2(MP2) calculations to these species as well as the fluoroacetylenes to ascertain the capability of these methods to furnish accurate enthalpies of formation in unsaturated C2 fluorocarbons.

CALCULATIONS

The calculations were performed using the GAUSSIAN code [9,10] on HP-PARisk, CRAY Y-MP, IBM R6000, SUN-Sparc and SGI Power-Challenge computers. Geometries were calculated at both the HF/6-31G(d) and MP2(FU)/6-31G(d) levels for ethylene, acetylene and the

eight fluorinated derivatives. Vibrational frequencies (scaled by 0.8929) were obtained using the HF/6-31G(d) basis. These quantities are required for determination of the ground state G2 and G2(MP2) electronic energies and for calculation of the temperature dependence of molecular heat capacities, entropies and reaction enthalpies. The G2 and G2(MP2) procedures have been presented in the original articles [11,12] as well as in an earlier paper from this laboratory [3].

Molecular and atomic electronic energies and heat capacities, together with elemental enthalpies of formation, were utilized to evaluate molecular enthalpies of formation at 298.15 K [(298)]. The computational procedures have been detailed elsewhere [3,13].

RESULTS AND DISCUSSION

Experimental and calculated enthalpies of formation are presented in Table 3. Recommended literature values, with their estimated uncertainties, are taken from the NIST compilation [7] for fluoroethylenes and from the JANAF tables for the fluoroacetylenes [13].

Calculated G2 and G2(MP2) enthalpies are given in columns 3 and 4 of Table 3; the quantities in parentheses represent deviations from the literature data. It is clear from the table that these deviations are quite large for both methods, as evidenced by the RMS errors for the fluoroethylenes which are more than three times greater than the experimental uncertainties. One sees similarly large errors in calculated enthalpies of the fluoroacetylenes, although uncertainties in the experimental results are unusually high.

Significantly, one finds from the large negative average deviations that they are not randomly distributed. The systematic nature of errors in the calculated G2 and G2(MP2) enthalpies of formation are well illustrated in Figure 2, where values of [(calc) - (expt)] are plotted versus the number of C-F bonds (n_{CF}) for both the fluoroethylenes (closed circles) and fluoroacetylenes (open circles). Although there is significant scatter in the data, which may partially be attributed to errors in the experimental enthalpies, one sees that, for both methods, the negative deviations increase approximately linearly with n_{CF} .

This behavior is quite similar to that observed in our recent investigations of fluoro-, chloro- and chlorofluoromethanes [3-4], in which the errors in *ab initio* enthalpies of formation

were linearly dependent upon the number of C-F and C-Cl bonds. In these studies, the quality of the calculated results were improved markedly by the application of linear Bond Additivity Corrections (BAC's) proportional to the number each type of carbon-halogen bond. For fluoromethanes, the use of a single BAC, Δ_{CF} , representing the error per C-F bond, reduced the RMS deviation from experiment to a value substantially less than the RMS uncertainties in the reported literature values. In a later investigation of predictions of in fluoroethanes [5], it was found that there is a "heavy atom" interaction between the second carbon the fluorine, which increased the errors due to the C-F bonds. Incorporating an interaction parameter, f_c , reduced RMS errors in this series to values comparable to the reported RMS experimental uncertainties.

It is straightforward to apply a BAC to the calculated *ab initio* enthalpies of formation, (calc), to obtain corrected values, (BAC), via the relation: $\Delta H_f^\circ(\text{BAC}) = \Delta H_f^\circ(\text{calc}) - \Delta_{CF}n_{CF}$. To test the transferability of BAC's between different series of halocarbons, we have chosen to utilize this equation with values of Δ_{CF} and determined for the saturated HFC's [4, 5, 14] to obtain corrected G2 and G2(MP2) enthalpies of formation in the fluoroethylenes and acetylenes. The results are displayed in the last two columns of Table 3.

It is immediately apparent that there is a dramatic improvement in the quality of the calculated enthalpies. RMS deviations have dropped by a factor of 2 in the acetylenes and by 4-5 in the ethylenes and, for both series, are below the uncertainties in the reported experimental data. Additionally, as indicated by the small residual average deviations in both series, almost all of the systematic error has been removed.

As noted by Melius in his development of the BAC-MP4 method [15,16], the heavy atom interaction parameter is expected to be dependent upon bond length (increasing for shorter bonds). Thus, one expects that errors in the BAC-corrected enthalpies could be reduced still further by reoptimizing for the multiply bonded systems studied here [17]. However, we find it quite satisfying that use of a BAC and interaction parameter derived purely from data on fluoromethanes and ethanes yields such excellent agreement between theory and experiment in the unsaturated fluoroethylenes and acetylenes.

Finally, we discuss briefly the enthalpy of formation of fluoroacetylene. As shown in Table

3, there is an extremely high estimated uncertainty (63 kJ/mol) for this enthalpy, resulting directly from the fact that there has been no experimental determination of ; the value reported in the literature [13] was obtained using average bond enthalpies. However, based upon our observation that BAC-corrected enthalpies of formation in all of the fluorocarbons studied to date have been accurate to within approximately 10 kJ/mol (and usually much better) for molecules with accurately measured experimental heats, we propose that the value, $[\Delta H_f^\circ(\text{H-C}\equiv\text{C-F}) = 114 \pm 10 \text{ kJ/mol}]$, furnishes a more reliable estimate, with reasonable error limits, than available from average bond enthalpy calculations.

REFERENCES

- [1] G. Baes, G. ANPI Mag. 112 (1992) 43.
- [2] W. Tsang and A. W. Miziolek, eds, Halon Replacements: Technology and Science; ACS Symposium Series 611, American Chemical Society, Washington, D. C. (1995).
- [3] R. J. Berry, D. R. Burgess, Jr., M. R. Nyden, M. R.; Zachariah and M. Schwartz, J. Phys. Chem. **99** (1995) 17145.
- [4] R. J. Berry, D. R. Burgess, Jr., M. R. Nyden, M. R.; Zachariah, C. F. Melius and M. Schwartz, J. Phys. Chem. **100** (1996) 7405.
- [5] R. J. Berry, C. J. Ehlers, D. R. Burgess, Jr.; M. R. Zachariah, M. R. Nyden and M. Schwartz, "Halon Thermochemistry: Ab Initio Calculations of the Enthalpies of Formation of Fluoroethanes," J. Molec. Struct. (Theochem), (Submitted)
- [6] R. J. Berry, C. J. Ehlers, D. R. Burgess, Jr., M. R. Zachariah and P. Marshall, "A Computational Study of the Reactions of Atomic Hydrogen with Fluoromethanes: Kinetics and Product Channels," Chem. Phys. Letters, (Submitted)
- [7] "Thermochemical and Chemical Data for Fluorinated Hydrocarbons," D. R. F. Burgess, Jr., M. R. Zachariah, W. Tsang and P. R. Westmoreland, National Institute of Standards and Technology, Technical Note 1412, U. S. Government Printing Office, Washington (1995).
- [8] Fluoroethylenes are formed in flames by thermal decomposition of fluoroethanes, activated decomposition of fluoroethyl radicals, and fluoromethyl radical recombination reactions (Ref. 7)
- [9] Gaussian 92, Revision F.4, Gaussian, Inc.: Pittsburgh, PA, 1992.
- [10] Gaussian 94, Revision A.1, Gaussian, Inc., Pittsburgh, PA, 1995.
- [11] L. A. Curtiss, K. Raghavachari, G. W. Trucks and J. A. Pople, J. Chem. Phys. **94** (1991) 7221.
- [12] L. A. Curtiss, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **98** (1993) 1293.

- [13] M. W. Chase, Jr., C. A. Davies, J. R. Downey, Jr., D. J. Frurip, R. A. McDonald and A. N. Syverud, JANAF Thermochemical Tables, 3rd. ed. J. Phys. Chem. Ref. Data, Suppl. 1, 14 (1985).
- [14] From Refs. [4,5], $\Delta_{CF}[G2] = -6.51$ kJ/bond, $fc[G2] = 1.24$; $\Delta_{CF}[G2(MP2)] = -7.98$ kJ/bond, $fc = 1.20$.
- [15] Melius, C. F. Thermochemistry of Hydrocarbon Intermediates in Combustion. Applications of the BAC-MP4 Method. In *Springer-Verlag DFVLR Lecture Notes*; Springer-Verlag: Berlin, 1990.
- [16] Melius, C. F. Thermochemical Modeling I. Application to Ignition and Combustion of Energetic Materials. *Chemistry and Physics of Energetic Materials*; S. N. Kluwer Academic: New York, 1992.
- [17] For example, if fc is reoptimized to minimize RMS errors in the G2 enthalpies of the fluoroethylenes, then the value of the parameter increases slightly (to $fc=1.37$) and the RMS and average errors are reduced to 4.1 and -0.2 kJ/mol, respectively.

Table 1. Enthalpies of Formation in Fluoroethanes.^a

A. Ab Initio Enthalpies

Species	Exp.	G2 ^b	G2(MP2) ^b	CBS-4 ^b	CBS-Q ^b
CH ₃ -CH ₃	-84.1 ±0.4 [20]	-86.0 (-1.9)	-83.1 (1.0)	-92.1 (-8.0)	-81.5 (2.6)
CH ₃ -CH ₂ F	-263.2 ±1.6 [21]	-279.7 (-16.5)	-279.5 (-16.3)	-277.7 (-14.5)	-274.1 (-10.9)
CH ₂ F-CH ₂ F	-433.9 ±11.8 [22]	-461.4 (-27.5)	-464.1 (-30.2)	-450.9 (-17.0)	-454.8 (-20.9)
CH ₃ -CHF ₂	-500.8 ±6.3 [21]	-516.4 (-15.6)	-518.3 (-17.5)	-508.5 (-7.7)	-510.0 (-9.2)
CH ₂ F-CHF ₂	-664.8 ±4.2 [23]	-687.1 (-22.3)	-691.8 (-27.0)	-669.4 (-4.6)	-679.7 (-14.9)
CH ₃ -CF ₃	-745.6 ±1.6 [21]	-772.1 (-26.5)	-775.4 (-29.8)	-760.0 (-14.4)	-764.4 (-18.8)
CHF ₂ -CHF ₂	-877.8 ±17.6 [22]	-906.7 (-28.9)	-913.3 (-35.5)	-875.2 (2.6)	-898.2 (-20.4)
CH ₂ F-CF ₃	-895.8 ±4.2 [21]	-934.2 (-38.4)	-940.2 (-44.4)	-911.9 (-16.1)	-925.0 (-29.2)
CHF ₂ -CF ₃	-1104.6 ±4.6 [21]	-1145.9 (-41.3)	-1153.9 (-49.3)	-1116.0 (-11.4)	-1135.2 (-30.6)
CF ₃ -CF ₃	-1342.7 ±6.3 [21]	-1383.7 (-41.0)	-1392.9 (-50.2)	-1348.4 (-5.7)	NA
RMS	±7.7	28.6	33.6	11.3	19.5
AVG	--	-26.0	-29.9	-9.7	-16.9

Table 1. (Cont'd.)

B. Enthalpies Calculated with Bond Additivity Corrections.

Species	G2 ^b [BAC]	G2(MP2) ^b [BAC]	CBS-4 ^b [BAC]	CBS-Q ^b [BAC]	BAC-MP4 ^b
CH ₃ -CH ₃	-86.0 (-1.9)	-83.1 (1.0)	-92.1 (-8.0)	-81.5 (2.6)	-86.94 (-2.8)
CH ₃ -CH ₂ F	-271.6 (-8.4)	-269.9 (-6.7)	-275.5 (-12.3)	-267.9 (-4.7)	-272.35 (-9.2)
CH ₂ F-CH ₂ F	-445.3 (-11.4)	-444.9 (-11.0)	-446.5 (-12.6)	-442.3 (-8.4)	-445.96 (-12.1)
CH ₃ -CHF ₂	-500.3 (0.5)	-499.1 (1.7)	-504.1 (-3.3)	-497.5 (3.3)	-505.34 (-4.5)
CH ₂ F-CHF ₂	-662.9 (1.9)	-663.1 (1.7)	-662.8 (2.0)	-661.0 (3.8)	-671.5 (-6.7)
CH ₃ -CF ₃	-747.9 (-2.3)	-746.7 (-1.1)	-753.4 (-7.8)	-745.7 (-0.1)	-755.44 (-9.8)
CHF ₂ -CHF ₂	-874.4 (3.4)	-875.0 (2.8)	-866.3 (11.5)	-873.2 (4.6)	-883.31 (-5.5)
CH ₂ F-CF ₃	-901.9 (-6.1)	-901.9 (-6.1)	-903.0 (-7.2)	-900.0 (-4.2)	-913.3 (-17.5)
CHF ₂ -CF ₃	-1105.5 (-0.9)	-1106.0 (-1.4)	-1104.9 (-0.3)	-1104.0 (0.6)	-1124.1 (-19.5)
CF ₃ -CF ₃	-1335.3 (7.4)	-1335.4 (7.3)	-1335.1 (7.6)	NA	-1357 (-14.3)
RMS	5.6	5.2	8.3	4.3	11.5
AVG	-1.8	-1.2	-3.0	-0.3	-10.2

a) $\Delta_f H^\circ$ at 298.15 K in units of kJ/mol.

b) Values in parentheses represent deviations from experiment.

Table 2. Bond Additivity Correction Parameters in Fluoroethanes.

Method	Δ_{CF}	f_c	RMS1 ^b	AVE1 ^b	RMS2 ^c	AVE2 ^c
Expt.			±4.1		±4.8	
A. Ab Initio						
G2	0.0	0.0	14.8	-13.1	28.6	-26.0
G2(MP2)	0.0	0.0	18.4	-15.7	33.6	-29.9
CBS-4	0.0	0.0	2.8	-1.7	11.3	-9.7
CBS-Q	0.0	0.0	8.6	-7.0	19.5	-16.7
B. Δ_{CF}						
BAC-G2	-6.51 ±0.41	0.0	3.2	-0.1	7.9	-6.5
BAC-G2(MP2)	-7.98 ±0.38	0.0	2.6	0.2	7.6	-6.0
BAC-CBS-4	-1.28 ±0.74	0.0	3.3	0.9	8.9	-5.8
BAC-CBS-Q	-3.51 ±0.55	0.0	1.6	0.1	9.4	-7.6
C. $\Delta_{CF} + f_c$						
BAC-G2	-6.51 ±0.41	1.24 ±0.08	3.2	-0.1	5.6	-1.8
BAC-G2(MP2)	-7.98 ±0.38	1.20 ±0.06	2.6	0.2	5.2	-1.2
BAC-CBS-4	-1.28 ±0.74	1.73 ±1.50	3.3	0.9	8.8	-3.0
BAC-CBS-Q	-3.51 ±0.55	1.78 ±0.29	1.6	0.1	4.3	-0.3

a) in units of kJ/mol
b) fluoromethanes (C1)
c) fluoroethanes (C2)

Table 3. Enthalpies of Formation in Fluoroethylenes and Acetylenes.^a

Species	expt ^{b,c}	G2 ^c	G2(MP2) ^c	G2 ^c [BAC]	G2(MP2) ^c [BAC]
CH ₂ =CH ₂	52.4 ± 0.8	53.4 (1.0)	55.5 (3.1)	53.4 (1.0)	55.5 (3.1)
CH ₂ =CHF	-140.1 ± 2.5	-146.1 (-6.0)	-146.5 (-6.4)	-138.0 (2.1)	-136.9 (3.2)
CHF=CHF[Z]	-297.1 ± 10.0	-315.0 (-17.9)	-318.0 (-20.9)	-298.9 (-1.8)	-298.8 (-1.7)
CHF=CHF[E]	-292.9 ± 10.0	-318.3 (-25.4)	-321.6 (-28.7)	-302.2 (-9.3)	-302.4 (-9.5)
CH ₂ =CF ₂	-336.4 ± 4.0	-359.0 (-22.6)	-361.4 (-25.0)	-342.9 (-6.5)	-342.2 (-5.8)
CHF=CF ₂	-491.0 ± 9.0	-512.4 (-21.4)	-517.7 (-26.7)	-488.2 (2.8)	-489.0 (2.0)
CF ₂ =CF ₂	-658.5 ± 2.9	-693.0 (-34.5)	-700.8 (-42.3)	-660.7 (-2.2)	-662.5 (-4.0)
RMS	6.7	21.3	25.2	4.6	4.9
AVG		-18.1	-21.0	-2.0	-1.8
CH≡CH	226.7 ± 0.8	234.6 (7.9)	236.8 (10.1)	234.6 (7.9)	236.8 (10.1)
CH≡CF	125.5 ± 63.0	106.2 (-19.3)	105.6 (-19.9)	114.3 (-11.2)	115.2 (-10.3)
CF≡CF	20.9 ± 21.0	0.5 (-20.4)	-3.1 (-24.0)	16.6 (-4.3)	16.0 (-4.9)
RMS	38.3	16.9	18.9	8.3	8.8
AVG		-10.6	-11.3	-2.5	-1.7

a) Enthalpies at 298.15 K and 1 atm in units of kJ/mol.

b) Experimental enthalpies taken from Refs. 7 (fluoroethylenes) and 13 (fluoroacetylenes).

c) Values in parentheses represent deviations from experiment.

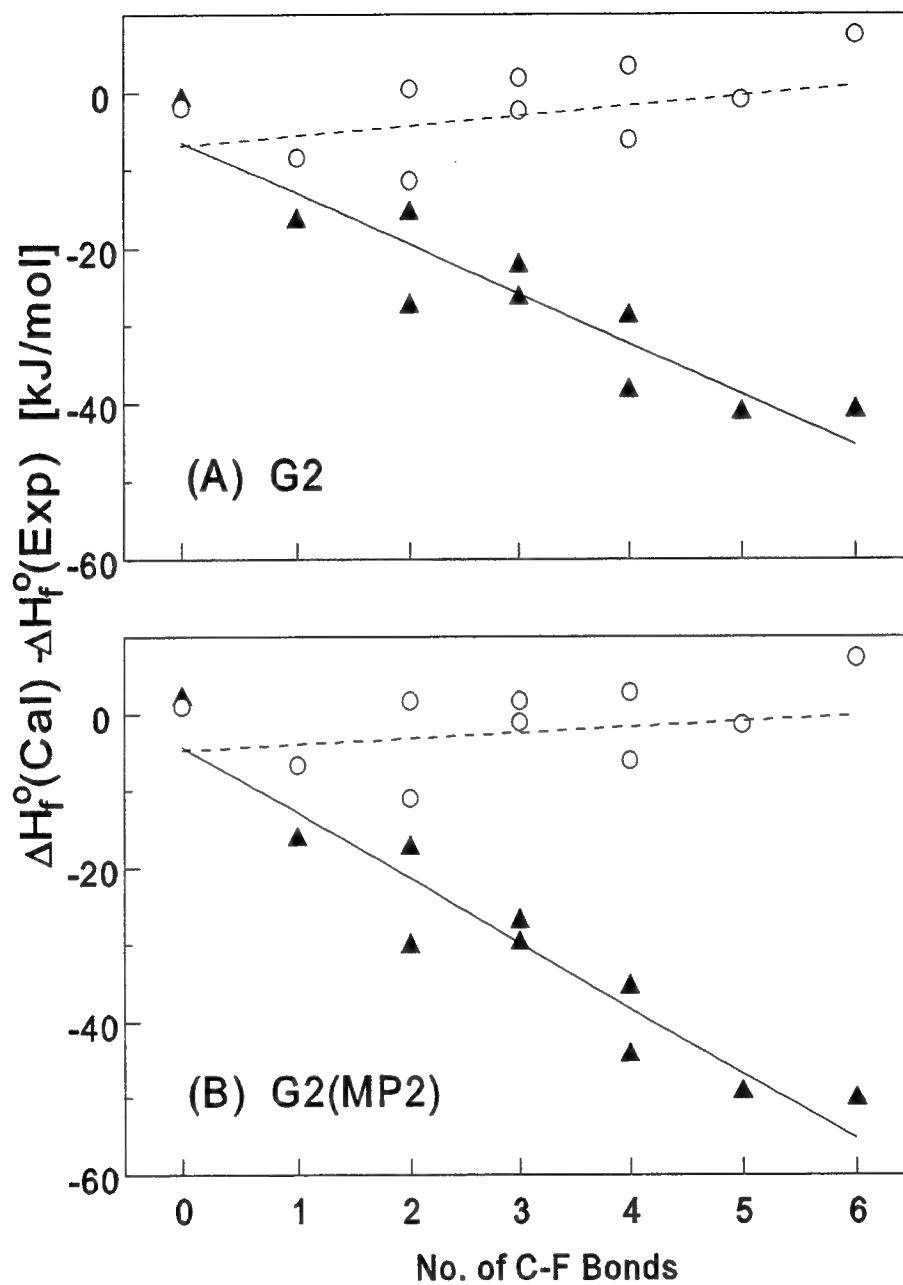


Figure 1. Errors in Enthalpies of Formation in Fluoroethanes. Closed symbols and solid lines - Ab Initio Enthalpies; Open Symbols and dashed lines - BAC corrected Enthalpies.

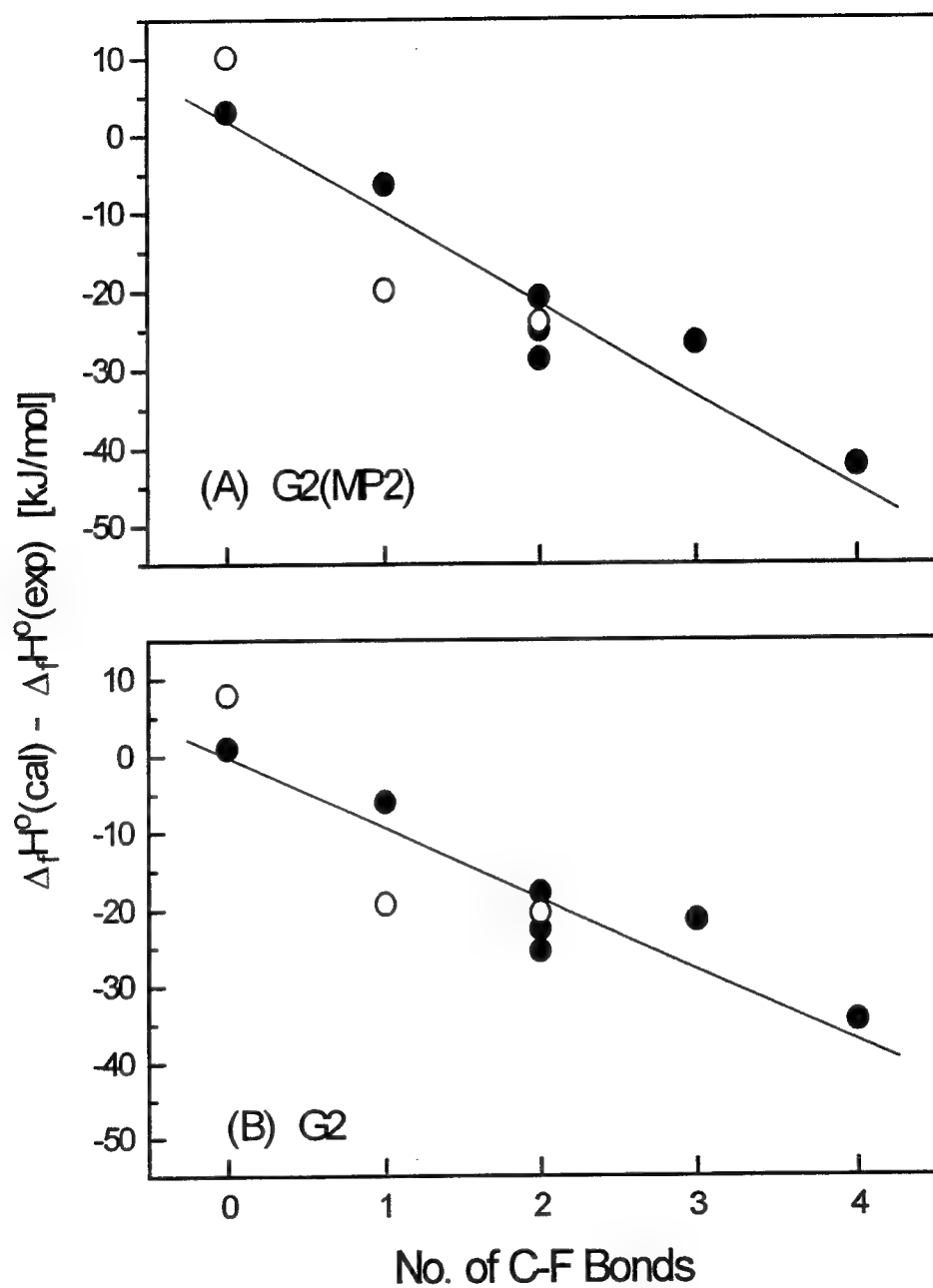


Figure 2. Errors in Enthalpies of Formation of Fluoroethylenes and Acetylenes. Closed symbols - fluoroethylenes; Open symbols - fluoroacetylenes.

A METHOD FOR STUDYING CHANGES IN TISSUE ENERGETICS
RESULTING FROM HYPERBARIC OXYGEN THERAPY

Thomas E. Skinner
Assistant Professor
Physics Department

Wright State University
3640 Colonel Glenn Hwy.
Dayton, OH 45435-0001

Final Report for:
Summer Faculty Research Program

Sponsored by:
Wright Laboratory
Wright-Patterson AFB, OH 45433
through the
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC

September 1996

A METHOD FOR STUDYING CHANGES IN TISSUE ENERGETICS RESULTING FROM HYPERBARIC OXYGEN THERAPY

Thomas E. Skinner
Assistant Professor
Physics Department
Wright State University

Abstract

This report describes development of a method for obtaining phosphorous (^{31}P) Nuclear Magnetic Resonance (NMR) spectra as a means for studying changes in tissue energetics resulting from hyperbaric oxygen (HBO) therapy. Since the method quantifies direct biochemical measures that are expected to be important in the process of wound healing, it is expected to provide a tool for assessing viability and metabolic activity levels of tissue. The methodology developed is capable of measuring levels of adenosine triphosphate (ATP), phosphocreatine (PCr), phosphomonoesters (PME), phosphodiester (PDE), and inorganic phosphate (Pi) in human tissue, including in and around wounds. Since the resonance position of inorganic phosphate depends on its state of protonation, ^{31}P NMR can also be used to determine intracellular pH, which is an additional important parameter in wound healing.

A METHOD FOR STUDYING CHANGES IN TISSUE ENERGETICS RESULTING FROM HYPERBARIC OXYGEN THERAPY

Thomas E. Skinner

Introduction

Recent advances in the physiology of tissue healing and wound repair have prompted the need to better understand the healing process. In particular, the impact of hyperbaric oxygen (HBO) therapy on the healing process and tissue energetics needs to be better understood. While HBO has proven to be advantageous in some wounds, decompression sickness, and gas emboli, the effects of HBO on the response to infection and the healing rates of lower extremity lesions has not been quantified. Correlation of treatment with the resulting success or failure of the wound healing process has been a cumbersome and somewhat subjective task. As a result, we proposed to investigate the feasibility of using phosphorous nuclear magnetic resonance (^{31}P NMR) spectroscopy to quantify changes in tissue energetics as a result of HBO treatment. We were subsequently invited by General Electric Medical Systems (GEMS) to participate in a clinical evaluation to determine the efficacy of a preproduction flexible ^{31}P surface coil designed to increase the quality of the detected NMR signal. The expertise of the Wright-Patterson USAF Medical Center Department of Hyperbaric Medicine was specifically cited in the invitation.

NMR spectroscopy is a noninvasive technique that uses static and time-dependent magnetic fields to excite emissions from nuclei such as hydrogen (^1H) and phosphorous (^{31}P) that have a magnetic dipole moment. The utility of NMR for investigating biological tissue lies in the fact that the compounds involved in energy metabolism and the major

metabolic pathways contain ^{31}P and ^1H , respectively. Since each compound resonates at a distinct frequency, changes in the strength of a given resonance reflect changes in the concentration of the corresponding metabolite. As a result, NMR can be used to provide a noninvasive "biopsy" of living tissue [1] and will allow the observation of the progression of expected biochemical and physiological changes that occur during tissue healing stages. The importance of ^{31}P for *in vivo* studies lies in the fact that it provides a window into the energy metabolism of living cells through the direct detection of adenosine triphosphate (ATP), phosphocreatine (PCr), phosphomonoesters (PME), phosphodiester (PDE), and inorganic phosphate (Pi) [2]. In addition, since the resonance position of inorganic phosphate depends on its state of protonation, ^{31}P NMR can be used to determine intracellular pH [3], which is an important parameter in wound healing. ^1H NMR provides access to an even larger group of chemical compounds such as fatty acids, carbohydrates, and free amino acids associated with the major metabolic pathways [4]. For example, changes in high energy phosphate levels, reflecting activity in the two primary energy-producing pathways: anaerobic glycolysis and the citric acid cycle. NMR spectroscopy is being used increasingly in metabolic studies of muscle and brain, two highly metabolic tissue types, but this study focused on assessment of tissues with vascular compromise, including normally less metabolically active tissue types.

A method that could access direct biochemical measures of wound healing would enhance the capabilities of the Hyperbaric Medicine Department by providing a means of both quantifying the wound healing process and simplifying the analysis of the wound's response to therapy. For the purposes of the present study, we hypothesized that wound healing as

a result of HBO treatment would be detectable through ^{31}P spectroscopy as changes in tissue energetics, long before the more visible signs of wound healing occurred. We therefore focused on devising a methodology for quantifying changes in energy metabolism by the direct measurement of the phosphorous compounds mentioned above. This methodology, dependent on both adequate hardware and software, is described in the next section. The potential of ^1H NMR for providing information on cell metabolism that might be important in wound healing was left for future study.

Methodology

Facilities at WPAFB Hospital relevant to the project include a 1.5 Tesla whole-body GE Signa Magnetic Resonance Imager/Spectrometer with state of the art hardware and software for data acquisition and analysis. However, the primary use of the Signa is for clinical imaging in the Hospital. Technical specifications for imaging are much less stringent than specifications required for spectroscopy. At the beginning of the summer, we discovered that the system did not "meet specs" for spectroscopy, and the service contract for the instrument does not include spectroscopy. A large effort was therefore required to characterize the system and determine the limits to obtaining useful ^{31}P spectra.

All spectra and images were acquired using standard protocols available on the Signa. An individual study begins with a localizer image to determine the precise anatomical region from which the spectrum will be obtained. The patient or volunteer lies down on the table assembly, which slides forward into the bore of the magnet. Prior to entering the magnet bore, the region to be imaged is landmarked using three perpendicular intersecting light beams, which register the anatomical region in the coordinate system used in the

magnet and determines how far the table slides forward such that the targeted region is located in the center of the magnet. The flexible ^{31}P coil is also placed around the anatomical region at this time. During the imaging procedure, the magnet homogeneity is optimized (*i.e.*, the magnet is shimmed) for the image acquisition. A standard fast spin-echo imaging sequence is used, requiring 16 seconds. The same shim parameters are then used when the spectrum is acquired using the GE program "fidcsi."

Results

To date, the primary limitation of the system is the homogeneity of the magnetic field, which affects signal linewidths and, hence, signal to noise ratio (SNR) of the ^{31}P spectra. (Note: the homogeneity is well-within specs for imaging, but less than optimal for spectroscopy).

We first determined that ^{31}P spectroscopy would not be feasible using the surface coils provided with the system. The low SNR of the system with these coils would require prohibitively long patient examination times. GE Medical Systems, as described in the Introduction, provided a flexible ^{31}P coil with significantly enhanced sensitivity compared to the surface coils. This increased sensitivity, as described below, is sufficient to provide useful spectra within the constraints of available clinical examination times. However, the full benefits of the flexible coil are still not being utilized, due to the aforementioned SNR limitations of the system.

A further major constraint on our work this summer has been attempting to do our research around the clinical schedule of the very busy MRI facility. The clinical imaging facility routinely maintains a patient backlog of 2-3 weeks or more, but will only consider

scheduling a research subject in the daytime (when HBO is available) if the MRI backlog is five days or less. We have thus done all our system testing and studies of normal subjects in the evenings and weekends. To date we have studied only one wound healing patient before and after HBO. That patient study was delayed about an hour post-HBO by clinical scheduling delays, probably reducing any HBO effects on tissue metabolic rates.

We characterized the performance of the flexible coil by acquiring 28 spectra from 6 healthy volunteers. Spectra were acquired from the calf, the ankle, and the ball of the foot—regions where wounds are typically seen in the patient population receiving HBO treatment. We have found that signal averaging 512 acquisitions provides sufficient SNR to quantify ATP, PCr, and Pi of non-muscle, non-brain tissue in healthy volunteers. The entire procedure requires 20-25 minutes. For the calf area, which is primarily muscle and provides higher signal intensity than the ankle or foot (*i.e.*, bone), only 64 acquisitions were necessary to obtain the same SNR in those volunteers.

After the spectrum is acquired, the data are transferred to the spectroscopy workstation (SUN 4/330) using the Unix operating system file transfer protocol (ftp). There, a spectroscopy analysis software package (SAGE) is used to perform the mathematical transformations necessary to quantify the data. The early weeks of the research effort were spent learning to use this package and assessing its capabilities. The software is well-designed and fairly straightforward to use, but is written to run on an older SUN operating system which is no longer standard and increased the learning curve. Nonetheless, we concluded that the software was adequate for the preliminary analysis of the ^{31}P spectra we were interested in performing.

Conclusions

The established methodology was applied to one patient at the end of the summer research program. The patient's wound was on the left foot. The spectrum acquired prior to HBO treatment showed significantly reduced (*i.e.*, almost absent) levels of PCr and Pi compared to the healthy control subjects. This is perhaps consistent with the anatomical and metabolic inhibitions to wound healing that characterize the patient population—in particular, arteriosclerotic vascular changes which occur in a substantial portion of the HBO patient population. After treatment, ATP levels increased slightly, but the increase was not significant compared to variations in the signal caused by the noise floor. As discussed above, the primary limitation of the system is the SNR, which is not within specifications. Nonetheless, we have devised and tested a viable means for monitoring changes in metabolites which are important markers for tissue energetics.

References

1. T. Lock, M. T. Abou-Saleh, and R. H. T. Edwards, *Br. J. Pshchiatry* **157** (suppl. 9), 38 (1990).
2. D. G. Gadian, G. K. Radda, R. E. Richards, and P. J. Seeley in *Biological Application of Magnetic Resonance*, R. G. Shulman, ed., Academic Press, New York, 463 (1979).
3. R. J. Gillies, J. R. Alger, J. A. Den Hollander, and R. G. Shulman, *Intracellular pH: Its Measurement, Regulation, and Utilization* in Cellular Functions, R. Nuccitelli and D. W. Deamer, eds., A. R. Liss, New York, 79 (1982).
4. J. W. Prichard and R. G. Shulman, *Ann. Rev. Neurosci.* **9**, 61 (1986).

INVESTIGATION OF STRUCTURAL DEFECTS IN 4H-SiC WAFERS

Marek Skowronski

Professor

Department of Materials Science and Engineering

Carnegie Mellon University

Pittsburgh, PA 15213

Final report for:

Summer Faculty Research Program

Wright Laboratory

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, Washington, DC

and

Wright Laboratory

September, 1996

INVESTIGATION OF STRUCTURAL DEFECTS IN 4H-SiC WAFERS

Marek Skowronski

Professor

Department of Materials Science and Engineering

Carnegie Mellon University

Abstract

Structural defects in silicon carbide wafers have been investigated using optical Nomarski-contrast microscopy, etching in molten KOH, and Auger spectroscopy. Several types of defects have been observed including: inclusions, low angle grain boundaries, micropipes, dislocations, cracks, and polishing scratches. Possible mechanisms responsible for formation of these defects have been proposed.

INVESTIGATION OF STRUCTURAL DEFECTS IN 4H-SiC WAFERS

Marek Skowronski

Introduction

Silicon carbide and especially its 4H polytype is a very promising semiconducting material for several important applications. Its high energy gap, high dielectric breakdown strength, high saturated electron velocity, and high thermal conductivity make it especially attractive for high temperature/high voltage/high power devices.^{1,2} However, there are number of unresolved problems that have to be addressed before this material could be reliably used in device processing. One of these problems is high density of extended defects in silicon carbide boules grown by Physical Vapor Transport method.³⁻⁶ The defect that attracted most attention so far is referred to as micropipes. This type of defect is a small diameter, typically one micron, empty tube that can penetrate most of the crystal length.⁷ Micropipes usually are aligned along c-axis which is also the most commonly used growth direction. These defects are known to reduce the breakdown voltages of SiC-based diodes resulting in formation of microplasmas.⁸ In addition to micropipes, SiC boules exhibit presence of platelet-like voids, dislocations, and low angle grain boundaries.

This project was focused on identifying the different types of extended defects present in 6H- and 4H-SiC wafers. Their densities and distribution has been determined and possible nucleation mechanisms proposed.

Experimental Procedures

All crystals examined in this project have been grown by Physical Vapor Transport. The crystals were cut into wafers with diameters in 1-2" range and finished with either one-side or two side polish. All wafers were oriented along basal plane with 3.5 to 8 ° miscut. The two-side polished wafers have been examined with Nomarski-contrast optical microscopy. This procedure highlighted low angle grain boundaries and strains associated with isolated micropipes. This method could not be used on one-side-polished samples and in particular on strongly absorbing n-type material. Even in cases where etching clearly revealed

presence of extended defects, optical contrast was nearly non-existent due to scattering on lapped side of the wafer and absorption on free carriers.

Optical inspection revealed presence of several types of inclusions in the volume of some wafers. The chemical make-up of one characteristic type described in detail below was analyzed using Auger electron spectroscopy. In order to avoid possible contamination caused by wafer processing (cutting, lapping and polishing) the samples were prepared by cleaving the wafer and analyzing the inclusions located on freshly cleaved surfaces.

Number of wafers have been chemically etched using molten KOH. The samples were degreased in TCE, acetone, and methanol followed by etching in KOH at 600 °C for 10-20 minutes. Different etch features on as-received Si-face were observed by Nomarski-contrast optical microscopy.

Results and Discussion

Several types of growth-related inclusions have been observed by optical microscopy on two-side polished wafers. One characteristic type of second phase inclusion is shown in Fig. 1 (a). Most inclusions of this type have a platelet-like shape with the lateral size in the basal plane being much larger than the thickness along c direction. Fig. 1 (a) presents the image with focus adjusted to the plane of inclusion. Typical sizes fall in the 50-300 μm range. Inclusions appear black with complex shape showing different degrees of hexagonal faceting. Another characteristic feature is the presence of narrow pipe-like structures at the periphery of most inclusions. Fig. 1 (b) shows the image of the same inclusion as in Fig. 1 (a) with focus adjusted to the top surface of the wafer. Four pinholes are apparent penetrating the wafer surface with sizes of less than 10 μm diameter. In most instances such pinholes have been observed on both sides of the wafer. It is likely that observed pinholes are the same type of a defect as micropipes, but the resemblance could also be superficial.

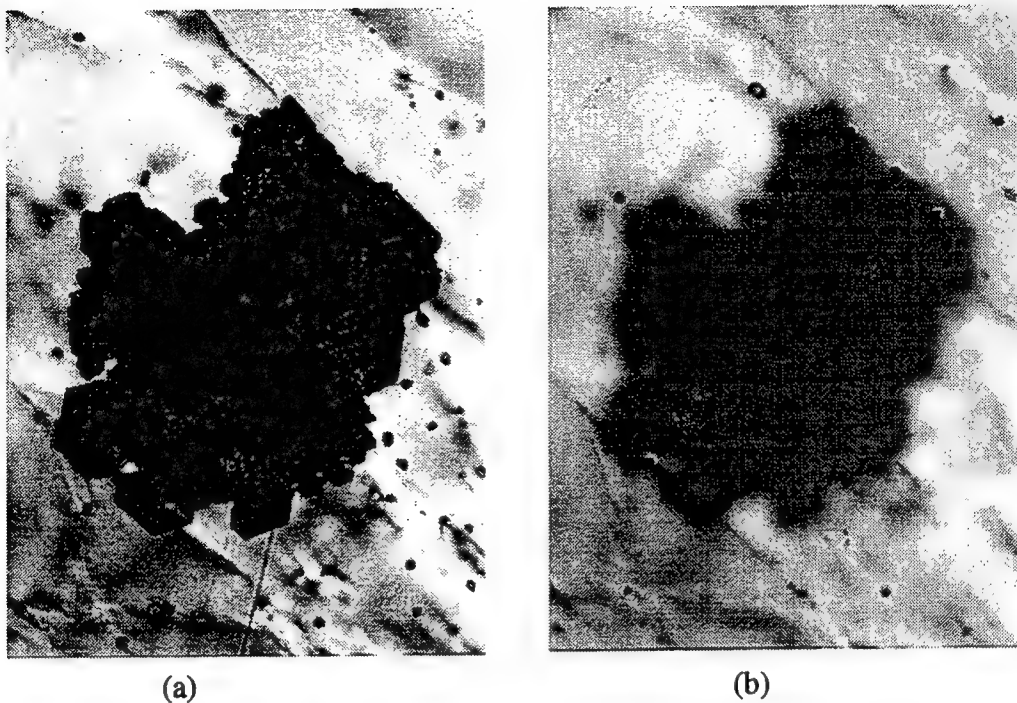


Fig. 1 Nomarski-contrast optical image of graphitic inclusion in SiC wafer, (a) focus on inclusion, (b) focus on the top surface of the wafer.

The origin of black inclusions has been investigated by scanning Auger electron spectroscopy. Total of four different inclusions from different boules have been analyzed. The optical image of cleaved surface of the 6H-SiC wafer with two inclusions marked A and B is shown in Fig. 2.

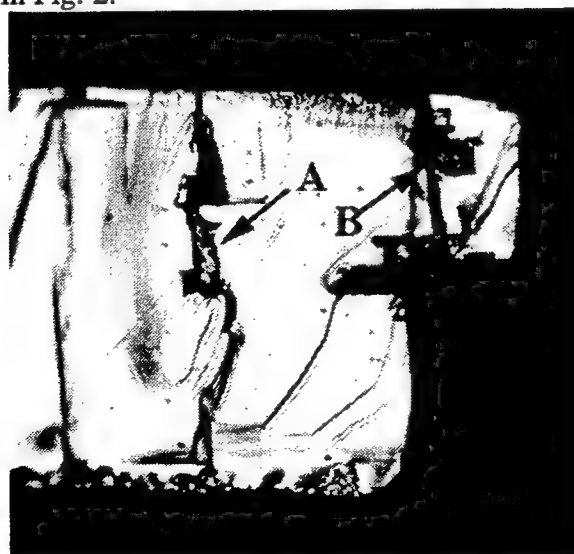


Fig. 2 Graphitic inclusions in SiC wafer cross section

Two Auger spectra obtained by focusing the beam on inclusion A and on the clean SiC matrix is shown in Fig. 3 (a) and (b) respectively. One difference is much higher ratio of

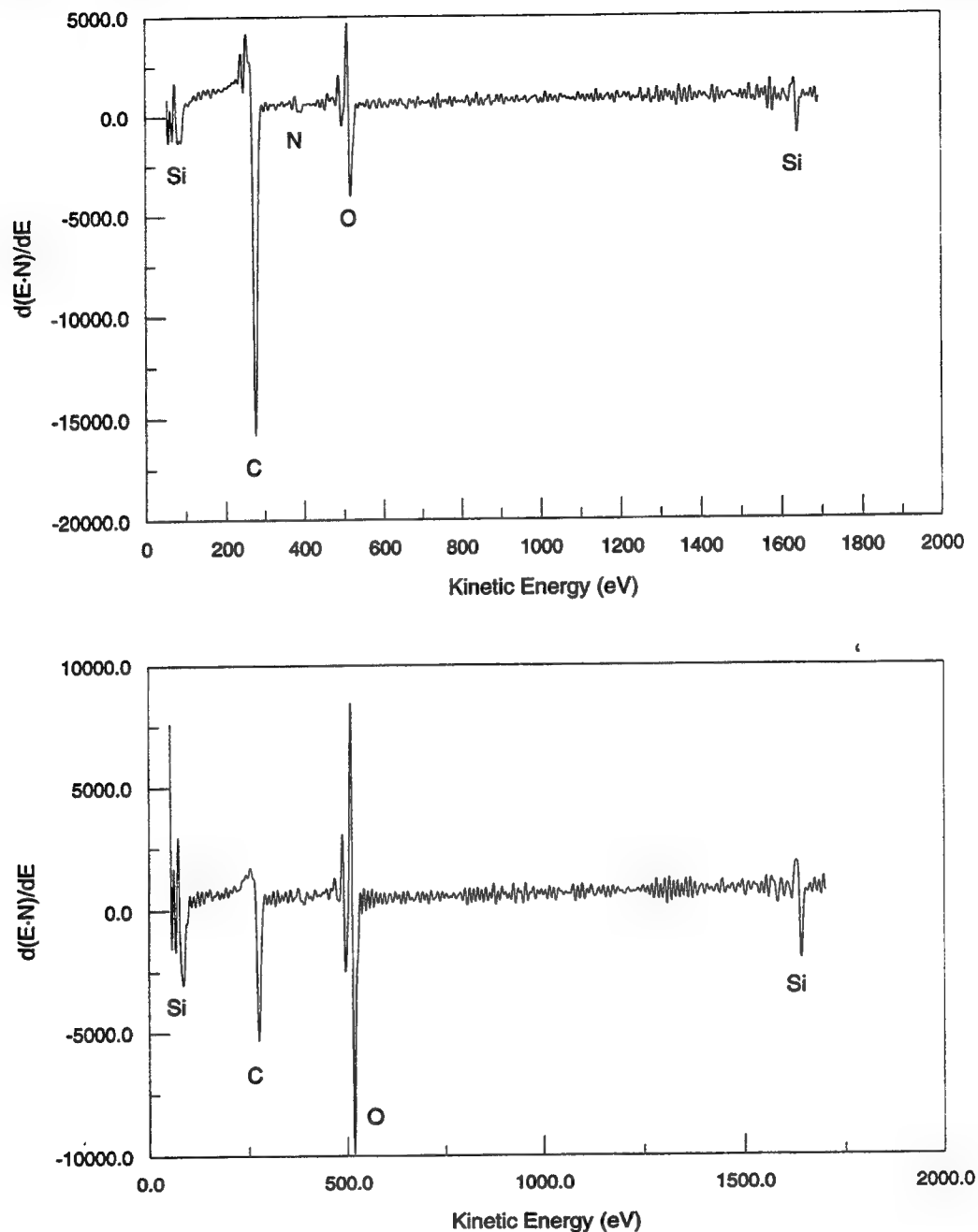


Fig. 3 Auger spectra obtained on the wafer cross-section from Fig. 2, (a) inclusion A (b) SiC matrix.

carbon/silicon peaks in the area of inclusion as compared to the matrix. This indicates that the inclusion is carbon rich.

In addition the shape of carbon peaks is different in both spectra. The expanded scale carbon peak obtained on inclusion A is shown in Fig. 4. Both the position of the main peak and shoulder located 262 eV is indicative of carbon-carbon bond.⁹ Based on these observations the black, faceted inclusions in SiC wafers have been interpreted as graphitic carbon.

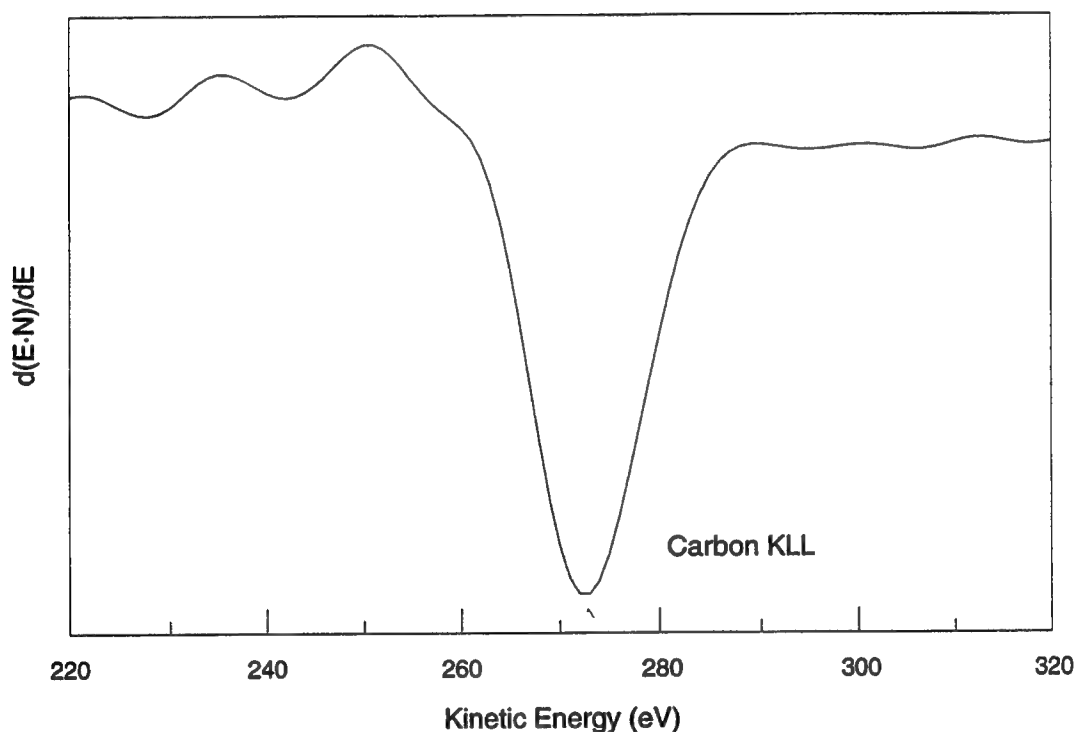


Fig. 4 Carbon peak in Auger spectrum obtained on inclusion A.

Carbon is an intrinsic element in the SiC growth system. The hot zone elements are made of high purity graphite and the SiC charge material contains carbon as well. One could expect formation of graphitic inclusions especially at the later stages of growth. Due to higher silicon vapor pressure and the fact that the hot zone is not leak tight, silicon is continuously removed from the hot zone at the higher rate than carbon. At the end of the run the supply of silicon from the charge material could no longer compensate the silicon loss from the seed. When this happens one should expect formation of carbon inclusions over the entire growth surface.

In a similar manner one can expect formation of silicon-rich conditions during SiC growth. The seed crystal which is a coldest part of the hot zone can serve a nucleation site for molten silicon evaporating from the charge material at early stages of growth. This mechanism can be used to explain the observation of second type of inclusions found in SiC wafers (Fig. 5). The inclusion is in the shape of a thin platelet parallel to the basal plane with well defined hexagonal outline. Typical size is in the 50-500 μm range. The volume of the defect is not completely filled with the second phase material. There appears to be only a thin layer of a second phase on the bottom of the platelet with most of the platelet volume remaining empty. One way of such defect being formed is if the growth of SiC was locally stopped or slowed down by formation of a second phase layer and after some tens of microns of growth the depression was overgrown from the sides encapsulating the second phase material and the void. Many A-type platelets show signs of liquid being present during crystal growth. Part of the thin colored layer at the bottom of the void in Fig. 6 appears to have formed droplets under the influence of surface tension. At this point we do not have a definitive proof of silicon inclusions being formed in the SiC crystals. Such evidence could be obtained however by using Auger electron spectroscopy and/or x-ray diffraction.

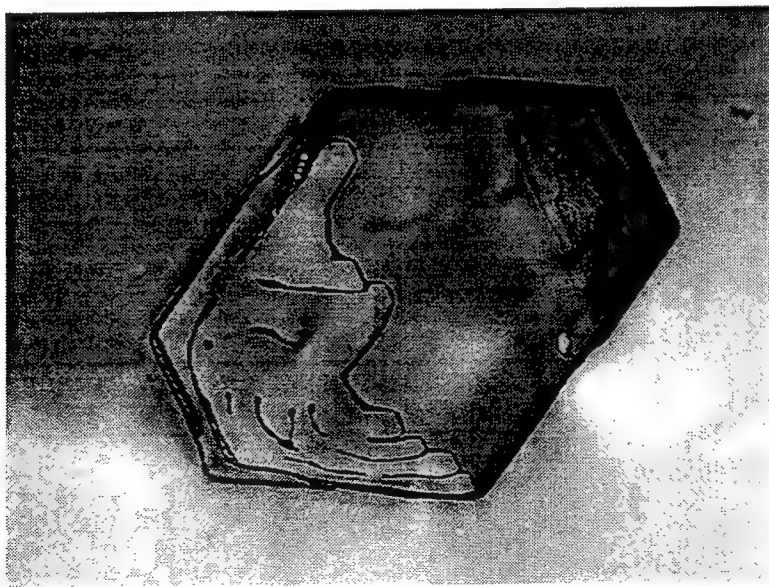


Fig. 5 Liquid-like inclusion in SiC wafer.

2. Dislocations

Dislocations in SiC wafers have been revealed by etching in molten KOH at 600 °C. Images of typical etch pits produced by etching are shown in Fig. 6. In n-type material the etch pits tend to be rounded with easily visible point at the bottom. Typical size ranges from 5 to 20 microns diameter. Bigger pits (such as one in upper left corner of Fig. 7 and second at the bottom of this figure) show some degree of faceting. In compensated SiC pits tend to form well defined hexagons. The difference between defects responsible for formation of different pits is not known at present. One of the possible explanations is the model involving dislocations of the same type (for example pure screw dislocations with Burger's vector along $\langle 0001 \rangle$) but different magnitude. Existence of such dislocations has been confirmed by Atomic Force Microscopy. Typical density of dislocations in SiC wafers is in the 10^3 - 10^5 cm⁻² range.

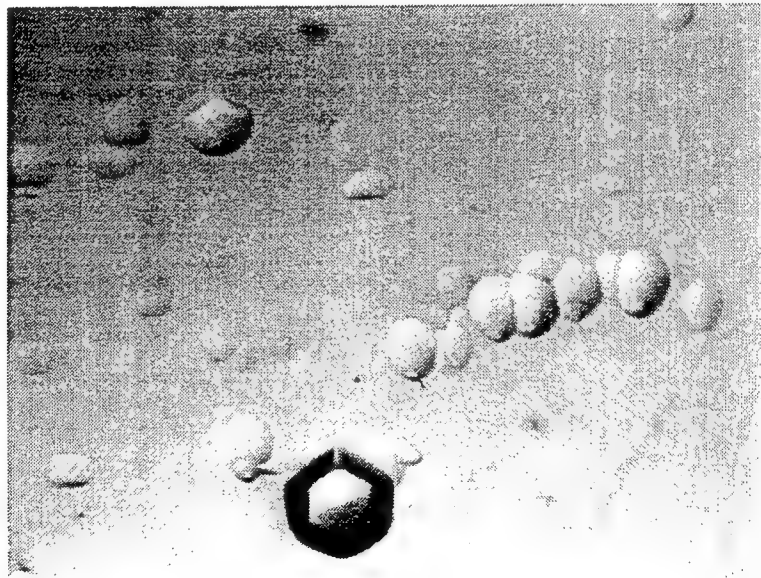


Fig. 6 Etch pits formed on SiC wafer after 15 minutes of etching in KOH at 600 °C.

Frequently, dislocation etch pits form well defined lines (Fig. 7). In some cases one can resolve individual pits along the line (for example the distance between dislocations in Fig. 7 is approximately 0.7 μ m). The lines are usually straight oriented along (10-10) directions. In some wafers the lines form easily recognizable hexagonal pattern. At this point there is no evidence of the nature of these dislocation arrays. However, their distribution suggests that they could be the result of dislocations gliding along the slip

planes. The source of stress at high temperatures which would be the cause of such deformation is not clear at present. The nature of dislocations constituting the lines could be investigated by Transmission Electron Microscopy. This, together with high resolution x-ray diffraction and x-ray topography of slip lines, will lead to positive identification of the mechanism leading to formation of the slip lines. It is always the first step in a sequence of events leading to elimination of defects and improvement of wafer quality.



Fig. 7 Dislocation arrays aligned along $\{1-100\}$

3. Micropipes

In addition to etch pits due to individual dislocations, another type of etch feature was observed (Fig. 8). It is usually well defined hexagons 2-4 times larger than dislocation etch pits. These features are commonly interpreted as due to micropipes. Typical density of micropipes in investigated wafers is in the $100-1000 \text{ cm}^{-2}$ range. Micropipes are not distributed uniformly throughout the wafer. Virtually always they appear in areas of high dislocation densities, most often along slip lines and grain boundaries. This association is in agreement with interpretation of micropipes as superdislocation formed by coalescence of individual dislocations. In effect micropipes serve as sinks for regular dislocations.

In addition to defects described above, wafers revealed presence of polishing-related scratches, cracks caused by stresses at low temperatures, and radial low angle grain boundaries. All of the above are detrimental to quality of epitaxial layers deposited on SiC wafers and performance of electronic devices.

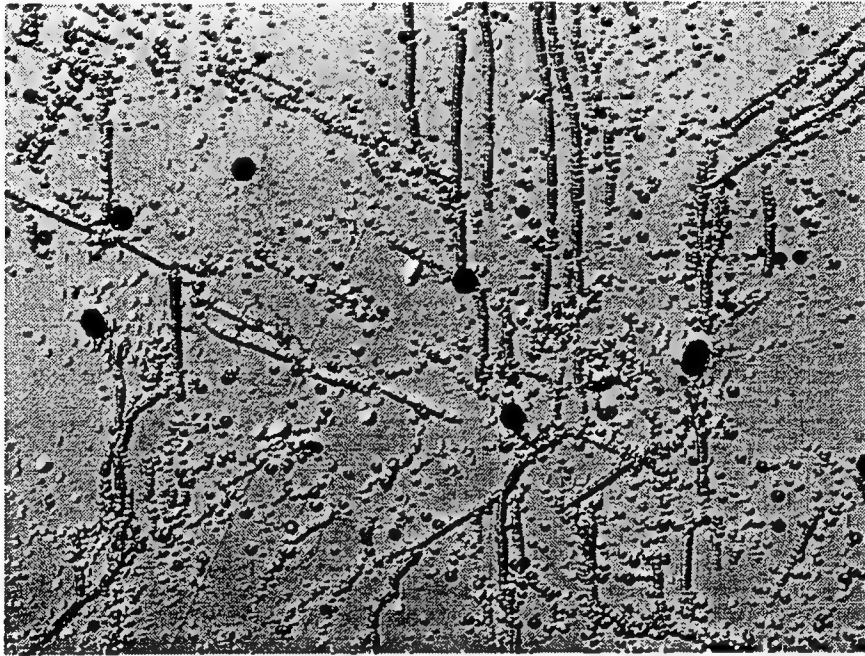


Fig. 8 Micropipes (large hexagonal pits) forming in the area of high dislocation density.

Conclusions

Extended defects in silicon carbide wafers grown by Physical Vapor Transport have been investigated. Two types of second phase inclusions have been observed. One of them was identified as due to graphitic carbon deposited on the growth surface at the late stages of SiC growth. The second type could be due to silicon droplets at the initial stages of growth. In addition to uniformly distributed background dislocations, a linear arrays have been observed. These arrays appear to be due to slip at high temperatures.

References:

- ¹R.F. Davis, G. Kelner, M. Shur, J.W. Palmour and J.A. Edmond, IEEE Proc. **79**, 677-701 (1991).
- ²H. Morkoc, S. Strite, G.B. Gao, M.F. Lin, B. Sverdlov and M. Burns, J. Appl. Phys. **76**, 1363 (1994).
- ³R.A. Stein, Physica B **185**, 211-216 (1993).
- ⁴R.C. Glass, L.O. Kjellberg, V.F. Tsvetkov, J.E. Sundgren and E. Janzen, J. Crystal Growth **132**, 504-512 (1993).
- ⁵M. Tuominen, R. Yakimova, R.C. Glass, T. Tuomi and E. Janzen, J. Crystal Growth **144**, 267-276 (1994).
- ⁶M. Dudley, S. Wang, W. Huang, C.H. Carter, V.T. Tsvetkov and C. Fazi, J. Phys. D: Appl. Phys. **28**, A63-A68 (1995).
- ⁷H.M. Hobgood, D.L. Barrett, J.P. McHugh, R.C. Clarke, S. Sriram, A.A. Burk, J. Gregg, C.D. Brandt, R.H. Hopkins and W.J. Choyke, J. Crystal Growth **137**, 181 (1994).
- ⁸J.A. Powell, P.G. Neudeck, D.J. Larkin, J.W. Yang and P. Pirouz, Inst. Phys. Conf. Ser. **137**, 161-164 (1994).
- ⁹J.T. Grant and T.W. Haas, Phys. Lett. **33A**, 386-387 (1970).

THEORETICAL INVESTIGATION OF PHTHALOCYANINE DIMERS

Grant D. Smith
Assistant Professor
Department of Chemical Engineering

University of Missouri-Columbia
Columbia, MO 65211

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Wright/Patterson Air Force Base, OH

and

Wright Laboratory

August 1996

THEORETICAL INVESTIGATION OF PHTHALOCYANINE DIMERS

Grant D. Smith
Assistant Professor
Department of Chemical Engineering
University of Missouri-Columbia

Abstract

The feasibility of performing quantum chemistry calculations of the binding energy of phthalocyanine (Pc) dimers as a method for establishing a potential energy surface for calibrating molecular mechanics force fields has been investigated. Calculation of the Pc dimer energy in an offset face-to-face configuration (C_{2h}) with MULLIKEN, a quantum chemistry code designed for efficient SCF and MP2 calculations on workstation platforms, takes approximately 30 IBM RS600 590 CPU hours at the SCF/6-31G* level. Calculation of basis set superposition error, which was found to be important, at the same level takes approximately the same computational effort. Therefore, an SCF potential energy surface, reflecting steric repulsion and electrostatic effects but not dispersion effects, with 30 data points can be determined in approximately 1800 CPU hours. Such a surface can be used to accurately determine repulsion and electrostatic parameters for a molecular mechanics potential energy function. Calculations on a dimer of the smaller but chemically similar azoporphine molecule indicate that MP2/6-31G* energies can be determined with approximately the same amount of computational effort as the SCF energies for the Pc dimer. An energy surface using correlated methods is useful for calibrating dispersion parameters in molecular mechanics force fields.

THEORETICAL INVESTIGATION OF PHTHALOCYANINE DIMERS

Grant D. Smith
Assistant Professor
Department of Chemical Engineering
University of Missouri-Columbia

Introduction

Organic compounds with extended π -electron systems form an important class of candidates for applications as nonlinear optical materials because of their potential large nonlinearities, varied structural features, and ease of processability¹. Metalloporphines are particularly attractive because they have a centrosymmetric π potential which can be modified by varying the central metal atom. Phthalocyanine (Pc) compounds are porphine derivatives that are of particular interest for optical limiting applications. A free-base (no central metal atom) phthalocyanine monomer (H_2Pc) with D_{2h} symmetry is illustrated in Figure (1). Many of the more interesting optical adsorption properties of Pc compounds occur as the result of interactions between molecules. A H_2Pc dimer, $(H_2Pc)_2$, with face-to-face stacking in D_{2h} symmetry, is illustrated in Figure (2). It is of great interest to understand the nature of interactions between Pc molecules, and how the optical properties of supramolecular complexes, such as dimers, are related to both the particular Pc (identity of the central metal atom) and the structure of the complex.

Among the most powerful methods for investigated molecular geometries, electronic structure, and intermolecular interactions, are ab initio molecular orbital calculations. Ab initio molecular orbital calculations have been carried out on Pc monomers using small basis sets.² Recently, ab initio SCF and density functional theory (DFT) calculations have been performed on a series of MPc ($M=Pb, Cu, Sn, H_2$) monomers using basis sets up to 6-31G**³. The feasibility of performing ab initio calculations on Pc dimers, due to the large size of the Pc molecule, and what information can be gleaned from

such studies, is unclear. Resolving these questions constitutes the major component of this work.

Intermolecular Interactions

Intermolecular interactions can be considered to be the sum of repulsion, dispersion, and electrostatic effects. In terms of classical, atomistic potential energy functions, it is common to represent the intermolecular energy by an expression of the form

$$E_{\text{intermolecular}} = \sum_i \sum_j A_{ij} \exp(-B_{ij} r_{ij}) - C_{ij}/r_{ij}^6 + q_i q_j / r_{ij} \quad (1)$$

Here, r_{ij} is the interatomic separation, and A_{ij} , B_{ij} and C_{ij} are potential energy parameters which depend upon the identity of atoms i and j . Together with the partial atomic charges q , these parameters make up the "force field". The sum is over all intermolecular atom-atom interactions. The first term, representing repulsion, or steric, effects, results from overlap of electron clouds. These interactions are always repulsive. The second term represents London dispersion forces, and is always attractive. Electrostatic effects result from interaction of permanent electric moments in one molecule with those in another. These permanent electric moment can be represented as atomic point charges. If the permanent electric moments are sufficient large, they can induce electric moments in other molecules, resulting in polarization effects which are not accounted for by eq 1. For charge neutral molecules, electrostatic interactions can be either attractive or repulsive, depending upon the orientation of the molecules.

It is of course desirable to examine intermolecular energetics using ab initio methods, rather than using empirical methods such as eq (1), because of the inherent approximations in such a representation. However, due the large size of the Pc molecules, and the large basis sets and correlated methods which must be used to accurately represent dispersion effects, such calculations are not currently feasible for Pc dimers. For example, our calculations on benzene dimers⁴ indicate that a basis set at least as large 6-311G(2d,2p)

is necessary to obtain accurate representation of dispersion effects. For $(H_2Pc)_2$, this basis set would yield 2164 basis functions.

Although ab initio molecular orbital methods cannot currently be employed to predict the structure of Pc dimers, semi-empirical methods such as SCMO have been used with some success in predicting dimer excited states and optical adsorption properties.^{5,6} However, these methods cannot be used to accurately predict the low energy structures of Pc dimers. Perhaps the best method currently available for predicting Pc dimer properties is to use molecular mechanics techniques to determine preferred structures and semi-empirical techniques to investigate electronic properties. However, in order for molecular mechanics methods to be useful and reliable, an atomistic force field for Pc molecules must be developed and validated. The purpose of this work has been to investigate the feasibility of using SCF calculations on Pc dimers as a method for obtaining a data base of structure and energy information sufficient to develop an accurate force field for Pc dimers.

Ab Initio Calculations on Pc Dimers

As discussed above, it is not currently feasible to use ab initio molecular orbital methods to predict the structure of Pc dimers because of the difficulty in representing dispersion effects. However, SCF ab initio methods with relatively modest basis sets have been demonstrated to accurately represent intermolecular repulsion and electrostatic effects⁷. Therefore, we have developed the following strategy to establishing a force field for Pc molecules:

- (1) Perform SCF ab initio geometry optimizations on single Pc molecules to obtain structural information and force constants needed to develop the bonded (stretching, bending, torsion and out-of-plane bending) force field.

- (2) Perform SCF ab initio electrostatic potential calculations on single Pc molecules to obtain partial atomic charges.
- (3) Perform SCF ab initio calculations on Pc dimers, maintaining the fixed monomer geometry given in (1). The intermolecular interaction energy is determined as a function of horizontal and vertical displacement of one monomer relative to the other.
- (4) Determine repulsion parameters by fitting molecular mechanics intermolecular energies to those generated in (3).
- (5) For the much smaller azoporphine illustrated in Figure (3), repeat (3) including electron correlation effects (MP2) to investigate dispersion effects.
- (6) Compare predictions of dispersion interactions using existing force fields with those obtained in step (5).

This study has concentrated on steps 1-3. Progress on these steps is reported below. Initial work on (5) is also briefly considered.

Calculations

Even at the SCF level, calculations on Pc dimers using reasonable basis sets are extremely computationally demanding. The prime purpose of this study has been to determine the feasibility of step (3). In order to do this, step 1 and 2 were performed as necessary preliminary steps. Calculations were performed using the computational chemistry code MULLIKEN⁸. This code was developed at the IBM Almaden Research Center specifically for use on RISC workstations, uses improved algorithms for performing SCF and MP2 calculations on large systems. These preliminary calculations were performed on a single IBM RS6000 node using 120 MB of memory. Minimal scratch space was required. It is our intention to perform future calculations in parallel using MULLIKEN at the Maui High Performance Computer Center.

Results

Step (1)

An optimized geometry SCF/6-31G* geometry was obtained for H₂Pc maintaining D_{2h} symmetry. The initial geometry was a SCF/3-21G geometry. The SCF energy and computation time for the geometry optimization are given in Table 1. H₂Pc at the 6-31G* level has 596 basis functions. At the 3-21G level, the D_{2h} symmetry structure has a single imaginary frequency. The true minimum structure has a much lower C_{2v} symmetry.³ However, optimization at the DFT level yields a structure of D_{2h} symmetry³. Therefore, in order to minimize computational time by maintaining maximum symmetry, the SCF D_{2h} symmetry is utilized. Figure (1) shows the labeling of the atoms. The internal coordinates for the optimized geometry are given in the appendix. See the appendix for instructions on how to obtain the coordinates via ftp.

Step (2)

Electrostatic potential charges were determined at the SCF/6-31G* and SCF/6-311G** levels using the SCF/6-31G* geometry. These calculations took 3.3 and 8.75 CPU hours, respectively. The charges are given in the appendix and are available via ftp. The charges do not have the symmetry of the molecule, indicating some problems with the calculation. Developers of MULLIKEN at the IBM Almaden Research Center are looking into the problem. Differences between the 6-311G** and 6-31G* charges are small.

Step (3)

SCF/6-31G* single point energy calculations were performed for the (H₂Pc) dimer using the monomer geometry obtained as described in Step (1). The dimer has 1192 basis functions with this basis set. The dimer energy was determined as a function of vertical displacement of the molecules, Δv , and horizontal displacement, Δh . With $\Delta h = 0$, the

dimer maintains the D_{2h} symmetry of the monomer. Calculations were performed for $\Delta h = 0$, $\Delta v = 3$ to 6 \AA . Dimer energies are given in Table I. Also shown is the dimer binding energy ΔE , given by

$$\Delta E = E_{\text{dimer}} - 2 E_{\text{monomer}} \quad (2)$$

At the closest separation of 3 \AA , the interaction between the monomers is highly repulsive. As the separation is increased, the repulsion drops rapidly. Even with a 6 \AA separation, where steric repulsion effects will be small, the dimer remains unbound. This indicates that electrostatic interactions are repulsive in the D_{2h} symmetry, as was found for the D_{6h} benzene dimer (BENZ), and as would be expected from interactions of permanent electric moments in this geometry. This appears to indicate that (attractive) polarization effects do not dominate the electrostatic dimer interactions.

A lower symmetry C_{2h} dimer geometry was also investigated, with $\Delta h = 2 \text{ \AA}$, $\Delta v = 4 \text{ \AA}$. Results are given in Table I. This dimer structure is apparently weakly bound at the SCF level, indicating attractive electrostatic interactions. However, molecular mechanics calculations (eq 1) in this geometry indicate that while the interactions of the permanent charges are less unfavorable in this geometry than in the D_{2h} structure, they remain repulsive. In contrast to what was found for the D_{2h} , this result seemed to indicate important polarization effects.

Before reaching the conclusion that polarization effects are important, i.e., that the C_{2h} structure studied is actually bound at the SCF level, basis set superposition error (BSSE) was first investigated. This was approximated by using the counter-poise method, calculating the stabilization of a monomer in the presence of the ghost orbitals of the other molecule. Twice this stabilization energy constitutes the counter-poise estimation of the BSSE. The dimer binding is considered to be less attractive by the amount of the BSSE estimation. The BSSE and BSSE corrected binding energies for the $\Delta v = 4 \text{ \AA}$, $\Delta h = 0$ and

2 Å structures are given in Table I. The BSSE corrections is relatively large compared to the total interaction energy, and after BSSE correction, the C_{2h} structure is no longer bound at the SCF level.

While computationally demanding, these results show that SCF single point calculations can be performed on systems as large as the H_2Pc dimer (1200 basis functions). Using the lower symmetry C_{2h} structure as a guide, calculating 30 points on the dimer potential energy surface, a reasonable number for parameterization of the repulsive part of an atomistic force field, would require about 30 X (26 dimer + 33 BSSE) = 1800 CPU hours on a high-end (590 or 3CT) IBM workstation. MULLIKEN is parallelized to run on SP2 machines, so, for example, these calculations could be accomplished in 100 hours employing 18 nodes.

An important question, however, remains as to the accuracy of the 6-31G* SCF dimer binding energies, given the (unexpectedly) large BSSE corrections. In order to explore this question, we have performed selected SCF/D95* calculations on the Pc dimer. BSSE effects are largely due to deficiencies in the core basis set. The 6-31G*, while having a double zeta valence description, has only a single zeta core description. The somewhat larger (monomer = 636 basis functions) D95* basis set has a full double zeta description. Results of these calculations, using the SCF/6-31G* monomer geometry, are given in Table I. As expected, the BSSE is significantly smaller for D95* basis set. However, the BSSE corrected binding energy is nearly the same as that obtained with the smaller basis set, indicating the adequacy of the 6-31G* BSSE corrected energies.

Step (4)

The force field will be generated by fitting to the quantum chemistry energy surface once a sufficient number of points have been generated. We have developed code for this purpose.

Step (5)

Azoporphine is illustrated in Figure (3). Four phenyl rings in the H_2Pc molecule have be replaced by hydrogen atoms. For this molecule, the 6-31G* basis set yields 356 basis functions, or 712 for the dimer. Single point SCF energy calculations for the dimer take approximately 3 CPU hours. A single point MP2/6-31G* calculation has been performed for the D_{2h} symmetry structure. This calculation required 30.7 CPU hours, and indicates that an potential energy surface for azoporphine can be generated at the MP2 level with about the same amount of work as the SCF surface for Pc dimer.

Conclusions

This work has demonstrated that accurate, BSSE corrected SCF dimer for Pc dimers can be determined using MULLIKEN, and that generation of a sufficient number of points on the dimer energies surfaces to parameterize the electrostatic and repulsion parts of the nonbonded potential is feasible. A similar conclusion holds for azoporphine at the MP2 level.

References

- (1) N. Matsuzawa, M. Ata and D. Dixon, *J. Phys. Chem.* **1995**, *99*, 7698.
- (2) B. Szczepaniak and P. Bragiel, *Vacuum* **1995**, *46*, 465.
- (3) P.N. Day, Z. Wang and R. Patcher, *MRS Spring Meeting*, **1996**, San Francisco, CA.
- (4) R.L. Jaffe and G.D. Smith, "A Quantum Chemistry Study of Benzene Dimer", in press *J. Chem. Phys.*
- (5) N. Ishikawa, O. Ohno, Y. Kaizu and H. Kobayashi, *J. Phys. Chem.* **1992**, *96*, 8832.
- (6) N. Ishikawa, O. Ohno, Y. Kaizu *J. Phys. Chem.* **1993**, *97*, 1004.
- (7) G.D. Smith, R.L. Jaffe and H. Partridge, "A Quantum Chemistry Study of the Interactions of Li^+ , Cl^- and I^- ions with Model Ethers", submitted to *J. Phys. Chem*
- (8) Mulliken: A Computational Quantum Chemistry Program developed by J. E. Rice, H. Horn, B. H. Lengsfeld, A. D. McLean, J. T. Carter, E. S. Replogle, L. A. Barnes, S. A. Maluendes, G. C. Lie, M. Gutowski, W. E. Rudge, Stephan P. A. Sauer, R. Lindh, K. Andersson, T. S. Chevalier, P.O. Widmark, Djamal Bouzida, G. Pacansky, K. Singh, C. J. Gillan, P. Carnevali, William C. Swope, and B. Liu, Almaden Research Center, IBM Research Division, 650 Harry Road, San Jose, CA 95120-6099

Appendix

The Pc and azoporphine SCF/6-31G* monomer geometries are available on curie.

ml.wpa.fb.af.mil anonymous ftp in directory pc_dimer.

Bonds and Bends for H₂Pc from SCF/6-31G* Geometry

bond	1	5	1.44808	
bond	1	9	1.35946	
bond	1	19	1.30069	
bond	5	6	1.39340	
bond	5	11	1.39465	
bond	9	23	.99443	
bond	11	15	1.37446	
bond	11	25	1.07351	
bond	15	16	1.40778	
bond	15	29	1.07497	
bond	19	33	1.32079	
bond	33	37	1.46427	
bond	33	41	1.34052	
bond	37	39	1.38676	
bond	37	43	1.38641	
bond	43	47	1.38285	
bond	43	51	1.07388	
bond	47	49	1.39928	
bond	47	55	1.07523	
bend	5	1	9	1.85135
bend	5	1	19	2.19069
bend	1	5	6	1.87647
bend	1	5	11	2.29261
bend	9	1	19	2.24115
bend	1	9	2	1.96914
bend	1	9	23	2.15702
bend	1	19	33	2.16808
bend	6	5	11	2.11411
bend	5	11	15	2.05221
bend	5	11	25	2.10654
bend	15	11	25	2.12443
bend	11	15	16	2.11687
bend	11	15	29	2.08958
bend	16	15	29	2.07674
bend	19	33	37	2.12366
bend	19	33	41	2.22913
bend	37	33	41	1.93040
bend	33	37	39	1.84059
bend	33	37	43	2.32354
bend	33	41	35	1.88280
bend	36	40	46	2.32354
bend	39	37	43	2.11906
bend	37	43	47	2.04880
bend	37	43	51	2.11408
bend	47	43	51	2.12031

bend	43	47	49	2.11533
bend	43	47	55	2.08878
bend	49	47	55	2.07908

Table I
SCF Energies for H₂Pc and (H₂Pc)₂

Species	Δh (Å)	Δv (Å)	P.G.	CPU (hrs)	E, SCF (Hartrees)	ΔE	BSSE (kcal/mol)	ΔE_{corr}
6-31G*								
mon	---	---	D _{2h}	21.0 ^a	-1657.891573	---	---	---
dimer	0	3	D _{2h}	15.5	-3315.560603	139.65		
dimer	0	4	D _{2h} (C _{2v}) ^b	12.0 (18.0)	-3315.776832	4.24	6.18	10.42
dimer	0	5	D _{2h}	9.5	-3315.780896	1.41		
dimer	0	6	D _{2h}	8.8	-3315.781559	1.00		
dimer	2	4	C _{2h} (C _s)	25.7 (33.0)	-3315.783624	-0.30	5.23	4.93
D95*								
mon	---	---	D _{2h}	6.8	-1658.108061	---	---	---
dimer	0	4	D _{2h} (C _{2v})	28.3 (18.5)	-3316.204433	7.33	2.78	10.11

^aThis calculation was a geometry optimization. ^bNumbers in parentheses are for BSSE (ghost orbital) calculations.

Table II
SCF and MP2 Energies for Azoporphine

Species	Δh	Δv	P.G.	CPU	E SCF MP2	ΔE SCF Corrected	MP2
	(Å)	(Å)		(hrs)	(Hartrees)	(kcal/mol)	
6-31G*							
mon	---	---	D _{2h}	7.25 ^a	-1047.233913	---	---
				2.25 ^b	-1050.533080		
dimer	0	4	D _{2h} (C _{2v})	4.75	-2094.465497	1.46	4.95
				30.7	-2101.095198		-18.22

^aThis calculation was a geometry optimization. ^bCalculated using GAUSSIAN94.
^cNumbers in parentheses are for BSSE (ghost orbital) calculations.

Figure Captions

Figure 1. A free-base phthalocyanine monomer with D_{2h} symmetry.

Figure 2. A free-base phthalocyanine dimer with D_{2h} symmetry.

Figure 3. Azophorphine with D_{2h} symmetry.

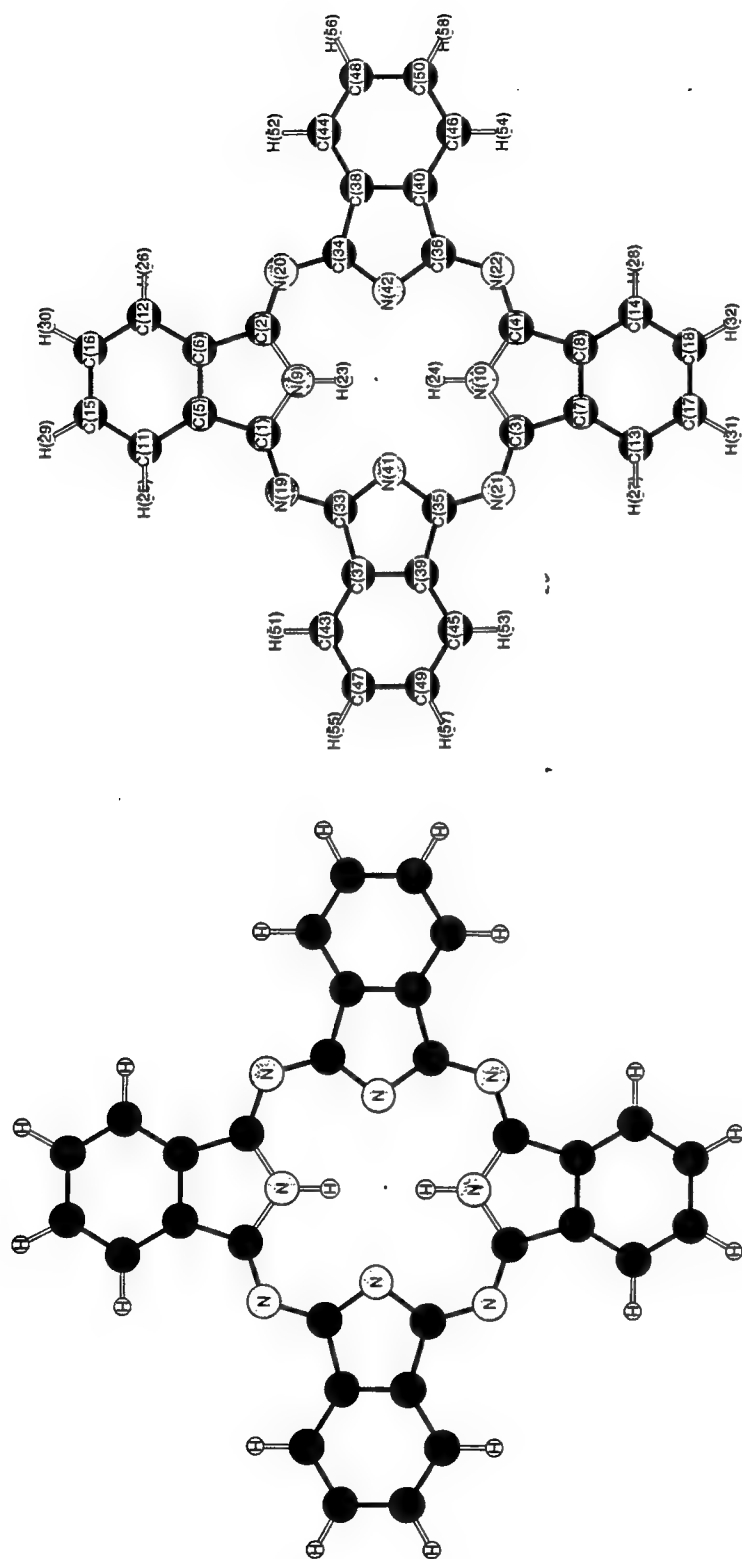


Figure (1)

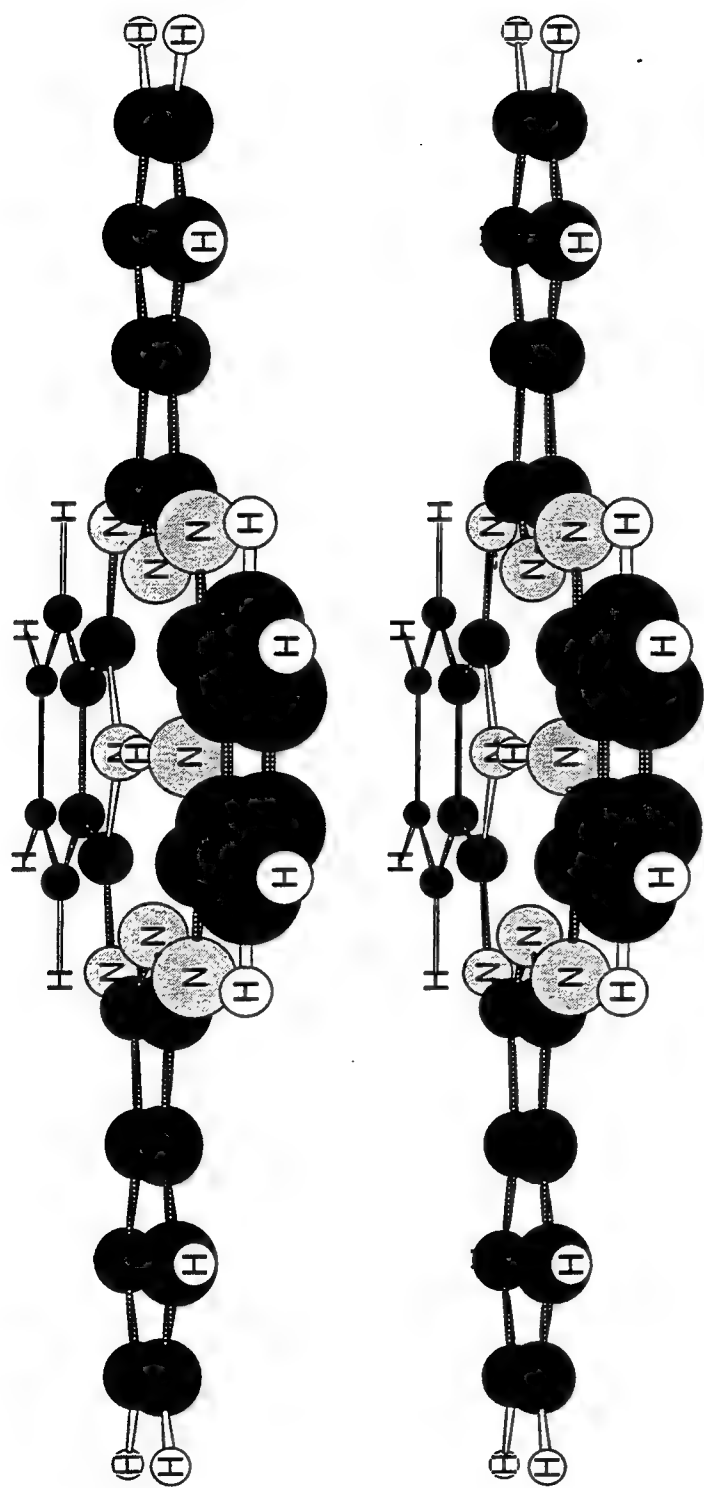
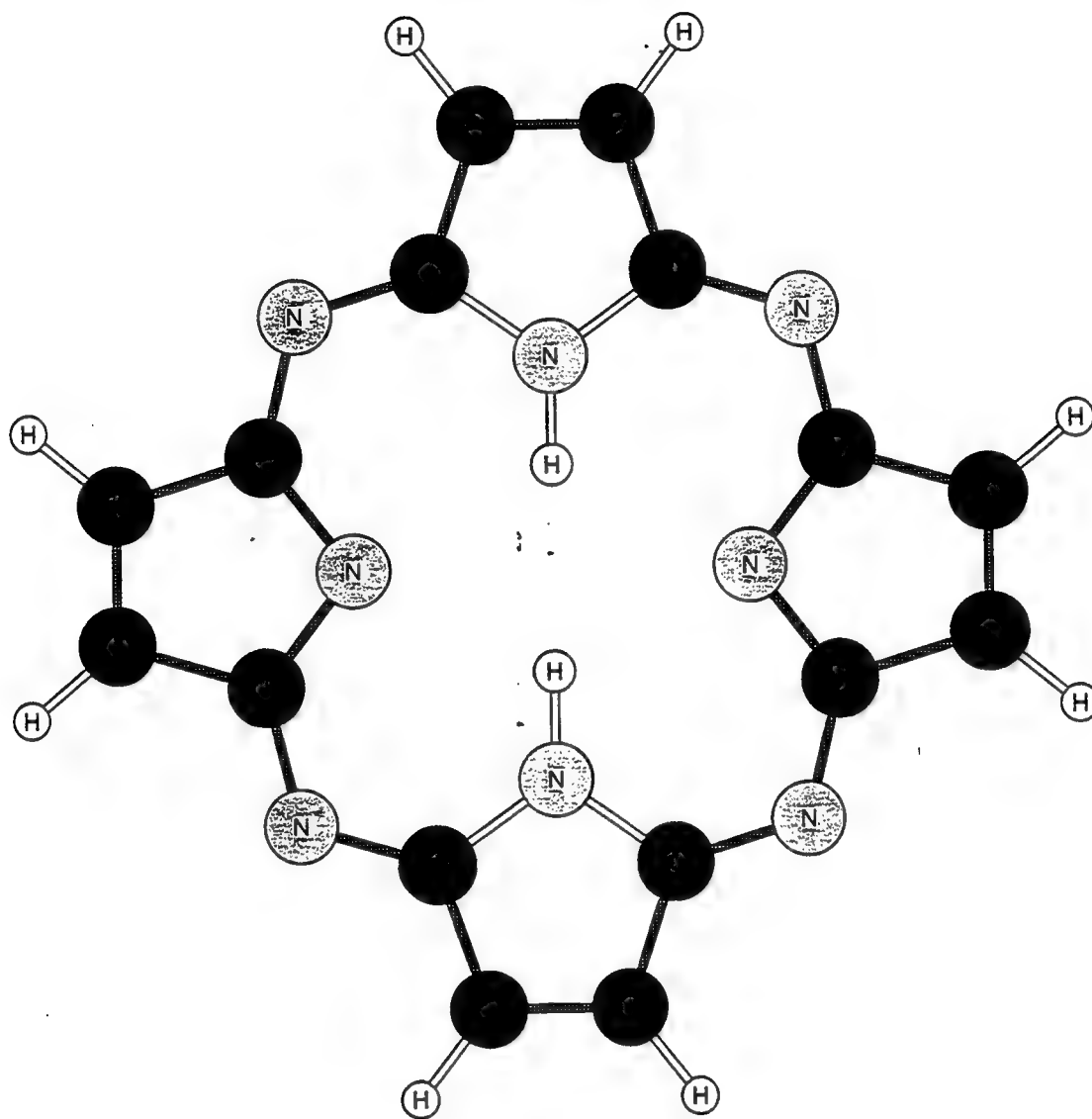


Figure (2)



55-19

Figure (3)

55-19

**Aging Aircraft: Preliminary Investigation
of Various Materials and Process Issues**

**James A. Snide
Professor
Graduate Materials Engineering**

**University of Dayton
300 College Park
Dayton, Ohio 45469-0240**

**Final Report for:
Summer Faculty Research Program
Wright Laboratory - Materials Directorate**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Wright Laboratory - Materials Directorate

August 1996

Aging Aircraft: Preliminary Investigation of Various Materials and Process Issues

James A. Snide
Professor
Graduate Materials Engineering
University of Dayton

Abstract

Initial effort during the AFOSR, eight-weeks, summer program at the Wright Laboratories was to try to identify some of the materials and process issues associated with aging Air Force aircraft. The approach was to review the current literature on aging aircraft, discussing the various topics with engineers from the Materials Directorate and attending the fourth Aging Aircraft Conference held at the USAF Academy. An annotated bibliography was prepared of the pertinent technical literature. Corrosion, nondestructive evaluation of corrosion damage and repair of corrosion damaged structures were identified as major problems. The research opportunities for the Air Force may be to improve inspection, corrosion inhibitions and control, characterization and analysis of aging structures and materials.

Reviewed by:

RL Crane
Dr. Robert Crane, WL/ML

Annotated Bibliography

1. Air Force 4th Aging Aircraft Conference, United States Air Force Academy, Colorado, 9-11 July 1996, jointly sponsored by Air Force Office of Scientific Research (AFOSR) and Wright Laboratory including the Flight Dynamics (WL/FI) and Materials Directorates (WL/ML).

The conference was held in conjunction with the Air Force Materiel Command (AFMC) Aging Aircraft Working Group Meetings. The purpose of the conference was to disseminate information on the status of R&D activities for the two working groups on Structural Integrity and Nondestructive Evaluation/Inspection. The papers on structural integrity were sub-classified as: widespread damage, corrosion/fatigue, dynamics and health monitoring and repairs. The papers on nondestructive evaluation were all presented under the categories of detection and multi-site damage detection. Although this conference presented papers of current research at universities, industry and government laboratories, the conference proceedings will probably not be available for nearly a year. The value of the conference was primarily to the participants although the program does list the people active in the field sponsored by the AFOSR, FAA and NASA.

2. FAA/NASA Symposium on Continued Airworthiness of Aircraft Structures -, hosted by FAA Center of Excellence for Computational Modeling of Aircraft Structures, Georgia Institute of Technology, Atlanta - 28-30 August 1996.

The themes of this symposium are to be the onset of WFD, role of NDI in aging aircraft, damage tolerance for commuter aircraft, residual strength predictions and engine integrity. This conference has less emphasis on corrosion than some of the other symposiums reviewed. A copy of the proceedings are to be mailed to the participants after the conference.

3. C.I. Chang and C. T. Sum, edited by, *Structural Integrity in Aging Aircraft*, 1995 ASME Mechanical Engineering Congress and Exposition, San Francisco, November 12-17, 1995.

This volume contains papers presented at the seven sessions of the Symposium on Structural Integrity in Aging Aircraft at the 1995 International Mechanical Engineering congress and Exposition. The topics of the papers cover a broad range of technical issues of aging aircraft, such as, multiple site damage, fatigue, corrosion, corrosion fatigue interaction, fretting fatigue, NDE for corrosion damage and repair techniques. The first article by Jim C.I. Chang of AFOSR presents an overview of the United States Air Force science and technology program on aging aircraft. The second article by John W. Lincoln presents the basic concept of the USAF Aircraft Structural Integrity Program and some of the technical challenges that were met as it was applied to USAF aircraft. This volume is probably the most current reference with published papers and reference with a broad coverage in the area of aging aircraft.

4. T. M. Cordell and R. D. Rempt, edited by, *Nondestructive Evaluation of Aging Aircraft, Airports, Aerospace Hardware and Materials*, Proceeding of APIE - the International Society for Optical Engineering, v. 2455, Bellingham, Washington, 6-8 June 1995.

5. C. Smith, *Revisiting the Aging Aircraft Nondestructive Inspection Validation Center - a Resource for the FAA and Industry*, Materials Evaluation, v. 53, p900-2, August 1995.

The Nation Aging Aircraft Research Program has established the Federal Aviation Administration (FAA) Aging Aircraft Nondestructive Inspection Validation Center (AANC) at Sandia National Laboratories in New Mexico. The AANC validated nondestructive inspection technology, provides a quick response capability, evaluates the reliability of NDI applications, and performs other projects to support the FAA and the aviation industry. Since it was formed, the scope of AANC activities have expanded into structural integrity analysis, repair evaluation and composite structure evaluation.

6. D. Roach, *Aircraft Inspection Validation Experiments and the Use of NDI Validation Samples*, Material Evaluation, v. 53, p803-7, July 1995.

Part of a special issue on nondestructive testing at the Aging Aircraft Nondestructive Inspection Validation Center. The important infrastructure – both hardware and procedural – employed by the center to validate nondestructive inspection methods for aircraft applications is discussed. The center's validation efforts will identify promising nondestructive inspection techniques and accelerate the technology transfer process to move these technologies into maintenance facilities.

7. J. Tirpak, *Washington Watch: The Aging of the Fleet*, Air Force Magazine, v. 79, n. 6, 1 July 1996.

The average age of USAF aircraft is rising steadily. With careful management, service leaders believe they can operate the oldest fleet ever safely and effectively.

8. C. E. Harris, *Overview of NASA Research Related to Aging Commercial Transport Fleet*, J. of Aircraft, v. 30 n. 1, p.64, 1 January 1993.

In an effort to ensure continued safe operation of the U.S. commercial airplane fleet, researchers are examining airframe and engine inspection and maintenance activities, as well as the structural designs themselves.

9. S. Derra, *Aging Airplanes: Can Research Make Them Safer?*, R&D 1990, v. 32, n1, p. 28-34, January 1990.

Inspection of older aircraft is explored. Boeing and McDonnell Douglas survey older aircraft and buy back some of the oldest for testing and analysis. Inspection techniques, including computer aided tomography, ultrasound and thermal imaging, are explored.

10. T. M. Cordell, *Life Management of Aging Air Force Aircraft: NDE Perspective*, Conference on Nondestructive Evaluation of Aging Infrastructure, pp. 34-44, Oakland, CA, June 1995.

Memory-Based Control Methodology with Application to EMRAAT Missile

David Y. D. Song
Assistant Professor
Department of Electrical Engineering

North Carolina A&T State University
1601 E. Market Street
Greensboro, NC 27411

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Airforce Base, DC

and

Wright Laboratory
Eglin Air Force Base

August 1996

Memory-Based Control Methodology with Application to EMRAAT Missile

David Y. D. Song
Assistant Professor
Department of Electrical Engineering
North Carolina A&T State University

Abstract

This work investigates a memory-based approach to controlling dynamic systems with significant nonlinearities and uncertainties. The main idea behind the memory-based method is to build the control scheme upon certain memorized information such as current system response, previous system response and past control experience. Fundamentally, the desired control signal in the scheme is "learned" and generated from observing and processing the most recent experience stored in a memory. System performance can be continuously improved during system operation. There is no need to repeatedly run the system for the same task (a process that is not allowed in many practical systems). Another advantage of this approach is that the overall required memory space does not grow with time and is much smaller than most existing learning control methods. The effectiveness of proposed method is verified via simulation on EMRAAT missile.

Memory-Based Control Methodology with Application to EMRAAT Missile

David Y. D. Song

I. Introduction

Since most practical systems are nonlinear in nature, how to deal with the inevitable system nonlinearities to achieve increasingly demanding design requirements has been the central theme of the current research, resulting in the development of various nonlinear control methodologies. It is interesting to note that in most existing control methods system nonlinearities are handled either by decomposing them into a linear combination of known nonlinear functions (as in nonlinear adaptive control) or by replacing nonlinear functions with multilayer static/dynamic neural networks or fuzzy systems.

In this research, we investigate a methodology referred to as *memory-based* approach for controlling nonlinear systems. The fundamental idea behind this approach is essentially different from traditional methods. More specifically, our control method *does*

- not assume the system is described by a linear model plus perturbations (as in H^∞ control);
- not attempt to linearize the system (as in nonlinear model-based or inverse dynamics control);
- not try to estimate certain parameters based on the linear parametric assumption (as in nonlinear adaptive control);
- not need to determine the bounds on certain nonlinear functions (as in robust control);
- not use infinite switch frequencies (as in variable structure control);
- not involve ad hoc membership functions (as in fuzzy control);
- not need to run the system repeatedly for the same task (as in repetitive learning control).

Instead, the control scheme is solely based upon certain memorized information such as current system response, previous system response and past control experience. Fundamentally, the desired control signal in the scheme is "learned" and generated from observing and processing the most recent experience stored in a memory. System performance can be continuously improved during system operation. There is no need to repeatedly run the system (a process that is not allowed in many practical systems). Another advantage of this approach is that the overall required memory space does not grow with time and is much smaller than most existing methods (Atkeson and Reinkensmeyer 1992, Schaal and Atkeson 1994).

Memory-based approach for solving engineering problems has a long history. Several researchers have contributed to this method. An early effort using memory-based concept was made by Fix and Hodges (1951) for a new pattern classification. Steinbuch and Piske (1963) implemented the direct storage of experience and nearest-neighbor

search process with a neural network for pattern recognition. Memory-based method has also been applied to weather prediction (Lorenz, 1969), speech recognition -- learning of pronunciation (Stanfill and Waltz, 1986), medical diagnosis, protein structure prediction (Waltz, 1987) and others (Atkeson and Reinkensmeyer 1992).

While memory-based concept may be applied to many areas, our special interest in this work lies in using this method for control synthesis of dynamic systems. To facilitate our discussion, we will concentrate on a class of systems described by

$$\dot{x}^{(n)} = [f_0(x) \pm \delta f(x)] + [g_0(x) \pm \delta g(x)]u \pm d(t) \quad (1)$$

where $x \in R$ is the output of interest (for instance, the position of a mechanical system), the scalar $u \in R$ is the control input and $x = [x, \dot{x} \dots x^{(n-1)}]^T \in R^n$ is the state vector of the system, $\delta f(\cdot) \in R^n$ denotes the uncertain nonlinearities involved in the system, $\delta g(\cdot) \in R$ represents the uncertainties in the control influence function, $g_0(\cdot) \in R$ denotes the nominal part of the control gain, f_0 is the known (nominal) part of the system nonlinearities (could be zero if no information is available) and $d(t)$ is a function representing the time profile of external disturbances. The functions $\delta f(x)$, $\delta g(x)$, and $d(t)$ are completely unpredictable both in magnitude and in direction, as usually the case in practice. To emphasize this fact we *exclusively* include the sign “ \pm ” in the model. In order for the system to admit a feasible control law, it must be assumed that the system is completely controllable for x in certain controllability region $R_c \subset R^n$ (Isidori 1989, Slotine and Li 1991). This can be ensured if $|g_0(x)| > |\delta g(x)| \forall x \in R_c \subset R^n$. We shall make this assumption in what follows.

The primary objective of this work is to develop a memory-based control scheme that is able to compensate modeling uncertainties, reject external disturbances and accommodate sub-system failures so that the state vector x follows a specified desired trajectory $x^* = [x^*, \dot{x}^* \dots x^{*(n-1)}]^T$. *Backward time-shift* method is used to develop memory-based control algorithms. This seems to be a logic and effective method to back up the memorized information needed in building the control scheme. Past experiences are processed by certain memory coefficients (MCs) which will be derived from system stability/performance consideration. The effectiveness of control algorithms are verified via simulation on EMRAAT missile.

II. Memory-based Controller

In this section, we discuss the basic structure of a memory-based controller and the fundamental issues associated with this type of controller. To begin, we introduce the following *information sets*:

$$\begin{aligned} S_u &= \{u_i | i = 0, 1, \dots, k-1 \text{ -- control history}\} \in R^k \\ S_x &= \{x_j | j = 0, 1, \dots, k-1 \text{ \& } k \text{ -- past \& current system responses}\} \in R^{k+1} \\ S_{x^*} &= \{x^*_q | q = 0, 1, \dots, k-1 \text{ \& } k \text{ -- past \& current desired system responses}\} \in R^{k+1} \end{aligned}$$

Here (and hereafter), u_k , x_k and x^*_k stand for $u(kT)$ -- control signal, $x(kT)$ -- state vector of the system, and $x^*(kT)$ -- the desired trajectory vector, respectively, and T is the sampling period. It is assumed that these *information sets* are stored in a content-addressable memory and are available for control design.

There could be many different ways to build a memory-based controller. One way is as follows,

$$u_k = \phi(z, w) \quad (2)$$

where $z \in R^h \subset S_u \cup S_x \cup S_{x^*}$ ($h \propto n, k$) is a vector associated with the stored system information, w denotes the memory coefficients (MCs) and $\phi(.) \in R^m$ is a (linear or nonlinear) function vector that maps z into u_k . While (2) is broad enough to cover a large class of memory-based learning controller (e.g., Arimoto et al 1985, Hara et al 1988, Atkeson and Reinkensmeyer 1992), it is worth noting that since it makes use of all the stored past information of the system, the required memory size may become extremely large as time goes by. This could prove disadvantageous in practice. Moreover, due to the large amount of information to be processed in (2), extensive computation and slow learning speed may be involved.

Understanding that both simplicity and effectiveness are important in control synthesis, we propose to build a memory-based control scheme by making use of the current and r -step past system responses and r -step past control history. That is, the controller is of the form

$$u_k = \phi(z_r, w_r) \quad (3)$$

where $z_r \in R^p \subset S_u^r \oplus S_x^r \oplus S_{x^*}^r$ ($p \propto n, r$). Here $1 \leq r \ll k$ is an integer representing the *order* of the controller and the subsets S_u^r , S_x^r and $S_{x^*}^r$ are defined as, respectively

$$S_u^r = \{u_i | i = k-r, \dots, k-1\} \subset S_u$$

$$S_x^r = \{x_j | j = k-r, \dots, k-1, k\} \subset S_x$$

$$S_{x^*}^r = \{x_{j^*} | j = k-r, \dots, k-1, k\} \subset S_{x^*}$$

The controller is conceptually illustrated as in Figure 1. The difference between (2) and (3) is that only the most recent experiences are honored in (3) while the past information beyond r -step is forgotten. This is motivated by the fact that for a practical system the latest system states have more influence on its future behavior.

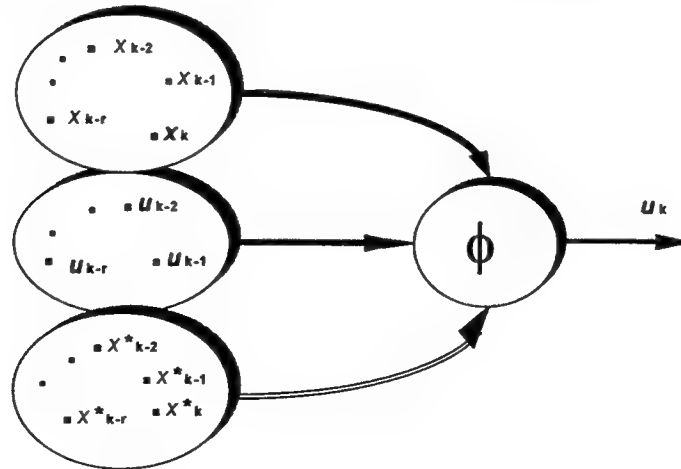


Figure 1. Constructing a controller using the nearest memorized information

A question that often arises with memory-based method is the effect of memory limitations (Schaal and Atkeson 1994). One of the favorable advantages of the proposed method is that the required memory size does not grow with time and is much smaller as compared with other methods. For example, to build a first order memory-based controller one only needs the following information:

$$S^1_u = \{u_{k-1}\}, S^1_x = \{x_k, x_{k-1}\} \text{ and } S^1_{x^*} = \{x^*_k, x^*_{k-1}\}$$

while constructing a second order memory-based controller only uses the stored information

$$S^2_u = \{u_{k-1}, u_{k-2}\}$$

$$S^2_x = \{x_k, x_{k-1}, x_{k-2}\}$$

and

$$S^2_{x^*} = \{x^*_k, x^*_{k-1}, x^*_{k-2}\}.$$

As can be seen, memory space is not an issue here.

Regarding the memory-based controller (3), if we view the memory-based controller as a *mechanism (network)* that *on-line* processes the memorized information and produces a control action at time instant $t = kT$, then design of a memory-based controller boils down to specifying such a mechanism, i.e., specifying the mapping function $\phi(\cdot)$ and the memory coefficients (MCs) w_i .

III. Preliminary Results

- For 1st-order system

$$\frac{dx}{dt} = f(x) + u = f_0(x) \pm \delta f(x, t) + u$$

The 1st-order memory-based controller is

$$u_k = w_0 u_{k-1} + w_1 x_{k-1} / T + w_2 x_k / T + w_{c_0} f_0(x_k) + w_{c_1} f_0(x_{k-1})$$

where the MCs are determined by backward time-shift method as follows,

$$w_0 = 1, \quad w_1 = 1, \quad w_2 = -2, \quad w_{c_0} = -1, \quad w_{c_1} = 1$$

The condition for the controller to apply is that for a sufficiently small T ,

$$|\Delta h_k| = |\delta f(x_k, t_k) - \delta f(x_{k-1}, t_{k-1})|_{t_k=kT} \leq \delta_1 < \infty$$

This condition is equivalent to that $h(x, t)$ is uniformly continuous with respect to x and t . As long as the system's dynamics do not change extremely fast, the above condition is always satisfied. It is interesting to note that the memory-based control is a nonlinear and dynamic compensator. In contrast to traditional dynamic control, however, the design procedure of the memory-based control is much simpler. Furthermore, the condition for the controller to apply is much weaker than other methods.

A sufficient condition that ensure the above relation is that the derivative of h w.r.t. time is bounded.

Similarly, the r^{th} -order memory-based controller can be constructed as follows

$$u_k = \sum_{i=1}^r w_{u_i} u_{k-i} + \sum_{i=0}^r w_{e_i} x_{k-i} / T + \sum_{i=0}^r w_{c_i} f_0(x_{k-i})$$

where the MCs w_{u_i} and w_{e_i} are determined as in the table and

$$w_{c_0} = -1$$

$$w_{c_i} = w_{u_i} \quad (\forall i = 1, \dots, r)$$

The condition for the controller to apply is

$$|\Delta^r h_k| = |\Delta^{r-1} h_k - \Delta^{r-1} h_{k-1}| \leq \delta_r < \infty$$

Determination of MCs

	w_{u_i}	w_{e_i}
$r=1$	1	-2 1
$r=2$	2 -1	-3 3 -1
$r=3$	3 -3 1	-4 6 -4 1
	4 -6 4 -1	-5 10 -10 5 -1
	5 -10 10 -5 1	-6 15 -20 15 -6 1
	6 -15 20 -15 6 -1	-7 21 -35 35 -21 7 -1
	7 -21 35 -35 21 -7 1	-8 28 -56 70 -56 28 -8 1
$r=8$	8 -28 56 -70 56 -28 8 -1	-9 36 -84 126 -126 84 -36 9 -1

$$r=1 \Rightarrow |e_k| = |x_k - x_k^*| \leq T^2 c_1$$

$$r=2 \Rightarrow |e_k| = |x_k - x_k^*| \leq T^3 c_2$$

- For n^{th} -order nonlinear systems with unit gain

$$x^{(n)} = f_0(x) + u \pm h(.)$$

The r -th order memory-based tracking controller is

$$u_k = \sum_{i=1}^r w_{u_i} u_{k-i} + \sum_{i=0}^r w_{e_i} s(e_{k-i}) / T + \sum_{i=0}^r w_{c_i} \eta_{k-i}$$

where

$$\eta(f_0, x^*) = f_0 + x^{*(n)} - \sum_{m=1}^{n-1} C_{n-1}^m q^{n-m} \gamma_0^m e$$

The condition for the controller to apply is

$$|\Delta^r h_k| = |\Delta^{r-1} h_k - \Delta^{r-1} h_{k-1}| \leq \delta_r < \infty$$

- For n^{th} -order nonlinear systems with constant non-unit gain

$$x^{(n)} = f_0(x) + g_0 u \pm h(.)$$

The r^{th} -memory-based tracking controller is

$$u_k = \sum_{i=1}^r w_{u_i} u_{k-i} + g_0^{-1} \left[\sum_{i=0}^r w_{e_i} s(e_{k-i}) / T + g_0^{-1} \sum_{i=0}^r w_{c_i} \eta_{k-i} \right]$$

where

$$\eta(f_0, x^*) = f_0 + x^{*(n)} - \sum_{m=1}^{n-1} C_{n-1}^m q^{n-m} \gamma_0^m e$$

The condition for the controller to apply is the same as before.

- For a more general nonlinear systems with time varying gain

$$x^{(n)} = f_0(x) + g_0(x) u \pm h(.)$$

where $h(.)$ represents the lumped uncertainties of the system. The r^{th} -order memory-based tracking controller is

$$u_k = g_{0_k}^{-1} \left\{ \sum_{i=1}^r w_{u_i} g_{0_{k-i}} u_{k-i} + \sum_{i=0}^r w_{e_i} s(e_{k-i}) / T + \sum_{i=0}^r w_{c_i} \eta_{k-i} \right\}$$

where

$$\eta(f_0, x^*) = f_0 + x^{*(n)} - \sum_{m=1}^{n-1} C_{n-1}^m q^{n-m} \gamma_0^m e$$

The condition for the controller to apply is the same as before.

Remark

1. Fundamentally the proposed memory-based controller does not presume slowly time-varying dynamics or periodic disturbances (Arimoto et al 1985). For the case of fast time-varying dynamics, a higher sampling frequency is needed to achieve better control precision. This seems to be understandable and typical of any control scheme because when system dynamics change faster, one must increase the speed in processing and updating the memorized information. Interestingly, since the controller has a very simple structure and involves simple arithmetic operations, its real-time computation demands very little CPU time, which makes it possible to choose a sampling frequency as high as desired.
2. It should be emphasized that in general $h(\cdot)$ is unknown and complicated. It is not feasible to re-organized $h(\cdot)$ into a known regressor matrix times a constant vector. It is also a painful task to estimate the possible bound on $h(\cdot)$. The proposed memory-based control avoids these difficulties. Moreover, there is no need to re-design or re-program the control scheme for minor change of environment.

IV. Application to EMRAAT Missile

The 2-D airframe of the EMRAAT missile is illustrated in Figure 2. Because of the asymmetric configuration, the missile can achieve bank-to-turn (BTT) maneuver. Such a configuration, however, leads to strong coupling effects between yaw, pitch and roll motions. The dynamic equations are given in Appendix. For technical discussion, we re-express these equations in the following matrix form,

$$J \begin{bmatrix} \dot{p} \\ \dot{q} \\ \dot{r} \end{bmatrix} + \begin{bmatrix} p \\ q \\ r \end{bmatrix} \times J \begin{bmatrix} p \\ q \\ r \end{bmatrix} = Q S d \left\{ \begin{bmatrix} n_1(\alpha, \beta) \\ n_2(\alpha, \beta) \\ n_3(\alpha, \beta) \end{bmatrix} + \begin{bmatrix} C_{l_{\delta_p}} & 0 & C_{l_{\delta_r}} \\ 0 & C_{m_{\delta_q}} & 0 \\ C_{n_{\delta_p}} & 0 & C_{n_{\delta_r}} \end{bmatrix} \begin{bmatrix} \delta_p \\ \delta_q \\ \delta_r \end{bmatrix} \right\}$$

where

$$J = \begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{xy} & I_{yy} & -I_{yz} \\ -I_{xz} & -I_{yz} & I_{zz} \end{bmatrix} = \begin{bmatrix} 1.08 & -0.274 & 0.704 \\ -0.274 & 70.13 & -0.017 \\ 0.704 & -0.017 & 70.66 \end{bmatrix} (EMRAAT)$$

$$\neq \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix} (Non-EMRAAT)$$

Define

$$\Theta = \begin{bmatrix} \phi \\ \theta \\ \psi \end{bmatrix}, \omega = \begin{bmatrix} p \\ q \\ r \end{bmatrix}, n = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}, u = \begin{bmatrix} \delta_p \\ \delta_q \\ \delta_r \end{bmatrix}, C = \begin{bmatrix} C_{l_{\delta_p}} & 0 & C_{l_{\delta_r}} \\ 0 & C_{m_{\delta_q}} & 0 \\ C_{n_{\delta_p}} & 0 & C_{n_{\delta_r}} \end{bmatrix}$$

Since

$$\omega = A(\theta, \phi)\dot{\Theta} \quad \text{and} \quad \dot{\omega} = \dot{A}\dot{\Theta} + A\ddot{\Theta}$$

The dynamic equation in $\phi - \theta - \psi$ is

$$\ddot{\Theta} = f(.) + g(.)u$$

where

$$f(.) = (JA)^{-1} \{ QSdn - J\dot{A}\dot{\Theta} - (A\dot{\Theta}) \times J(A\dot{\Theta}) \}$$

$$g(.) = (JA)^{-1} QSdC$$

Define the tracking error

$$e = \Theta - \Theta^*$$

To use the memory-based design procedure as described before, we introduce

$$s = \dot{e} + \gamma_0 e$$

Then

$$\dot{s} = \ddot{e} + \gamma_0 \dot{e} = f(.) + g(.)u + \gamma_0 \dot{e} - \ddot{\Theta}^*$$

Taking into account the parameter perturbation, we have

$$\dot{s} = f_0(.) + g_0(.)u + \gamma_0 \dot{e} - \ddot{\Theta}^* + h(.)$$

The first order memory-based control for the system is

$$u_k = g_{0_k}^{-1} [g_{0_{k-1}} u_{k-1} + w_1 s_k / T + w_2 s_{k-1} / T + w_{c_0} \eta_k + w_{c_1} \eta_{k-1}]$$

with

$$\eta_k = f_{0_k} - \ddot{\Theta}_k^* + \gamma_0 \dot{e}_k$$

The r^{th} -order memory-based control is

$$u_k = g_{0_k}^{-1} \left[\sum_{i=1}^r w_{u_i} g_{0_{k-i}}^{-1} u_{k-i} + \sum_{i=0}^r w_{e_i} s_{k-i} / T + \sum_{i=0}^r w_{c_i} \eta_{k-i} \right]$$

The detail proof of the above results will be provided in the forthcoming technical paper.

V. Simulation

To verify the effectiveness of the proposed control method, we carried out a series of simulations on EMRAAT missile. The parameters used for simulation are as in the Appendix. All the parameters are used as the normal values. To test the robustness, we

introduced 15% parameter perturbation. The control objective is to achieve quick tracking of the missile in yaw, pitch and roll. The desired trajectories are given as,

$$\phi = 40\pi / 180(1 - e^{-t})(\text{deg})$$

$$\theta = 30\pi / 180(1 - e^{-t})(\text{deg})$$

$$\psi = 10\pi / 180(1 - e^{-t})(\text{deg})$$

The sampling period is $T = 0.01$ (s). The control results are as in Figure 3.

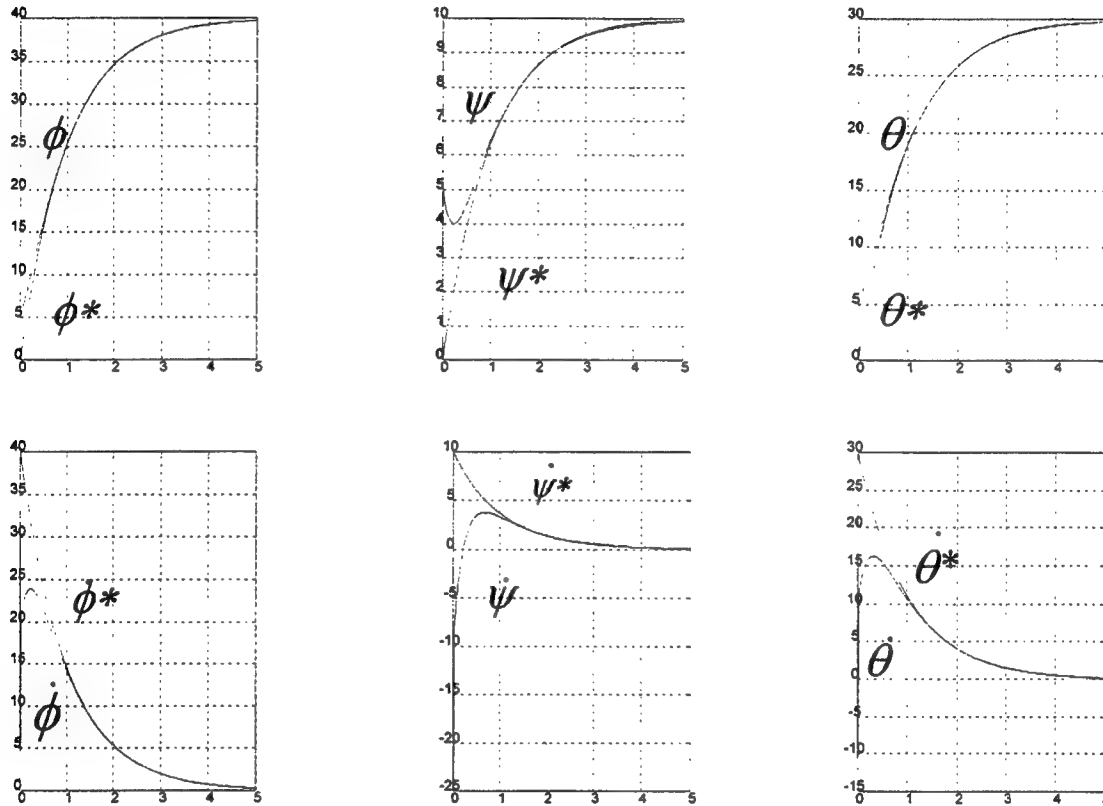


Figure 3. Yaw-pitch-roll motion control of EMRAAT missile

VI. Concluding Remarks

Our preliminary investigation indicates that the proposed memory-based control method exhibits attractive features. In fact, upon examining the memory-based control algorithms one may find that

1. The control scheme, purely built upon past control experience and most recent system responses, is quite effective in dealing with system nonlinearities, modeling uncertainties and external disturbances;
2. The proposed method demonstrates robust, adaptive and fault-tolerant properties, while these properties are achieved through an avenue that is essentially different from traditional methods -- there is no need for parameter estimation, repetitive

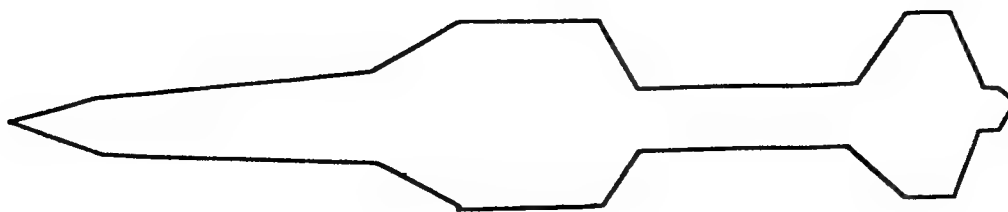
actions, infinite switching frequencies, ad hoc design process or extensive off-line training;

3. The controller is of digital form and is structurally simple and computationally inexpensive. These features make it particularly attractive for real-time implementation;
4. The novelty of the proposed method also lies in its flexibility in that the structure of the controller remains unchanged for different nonlinear functions $\delta f(.)$ and $\delta g(.)$. Furthermore, the memory size does not grow with time, which could significantly facilitate real-time implementation.

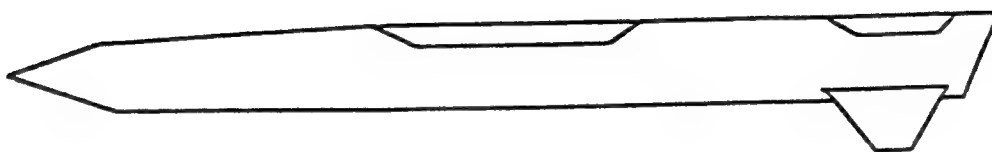
While memory-based approach for solving engineering problems has along history, applying this method to control systems is a new attempt. Our current results show that memory-based control has potential for control applications.

Acknowledgment

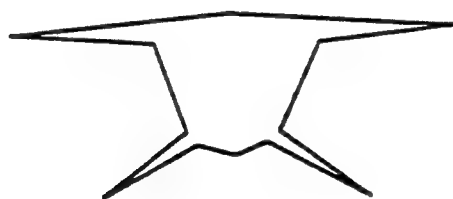
This work was partially supported by Air Force Office of Scientific Research under Summer Research Program. Technical guidance and assistance from Dr. James Cloutier is also gratefully acknowledged.



TOP VIEW



SIDE VIEW



REAR VIEW

Figure 2. EMRAAT Airframe

References

- [1] S. Arimoto, S. Kawamura, F. Miyazaki and S. Tamaki, "Learning Control Theory for Dynamic Systems," *IEEE Int. Conf. on Decision and Control*, 1985, pp.1375-1380.
- [2] C. G. Atkeson and D. J. Reinkensmeyer, "Using Associate Content-addressable Memories to Control Robots," *Neural Networks for Control*, Edited by W. T. Miller, R. S. Sutton, and P. J. Werbos, 1992, pp. 255-285.
- [3] A. G. Barto, C. V. Hollot and B. E. Ydstie, "Neural Networks for Adaptive Control", NSF Research Proposal, 1989.
- [4] F. C. Chen and C. C. Liu, "Adaptively Controlling Nonlinear Contiguous-Time Systems Using Multilayer Neural Networks," *IEEE Trans. on Auto. Control*, Vol. 39, No. 6, 1994, pp. 1306-1310.
- [5] F. C. Chen and H. K. Khalil, "Adaptive Control of a Class of Nonlinear Discrete-Time Systems Using Neural Networks," *IEEE Trans. Auto. Control*, Vol. 40, No. 5, 1995, pp. 791-801.
- [6] J. M. Elzebda, A. H. Nayfeh and D. T. Mook, "Development of an Analytical Model of Wing Rock for Slender Delta Wings," *J. of Aircraft*, Vol. 26, No. 8, 1989.
- [7] G. Cybenko, "Approximation by Superposition of a Sigmoidal Function," *Mathematics of Control, Signals, and Systems*, No. 2, 1989, pp. 303-314.
- [8] E. Fix and J. L. Hodges, Jr., "Discriminatory Analysis, Non-parameteric Regression: Consistency Properties," *Technical Report No. 4, USAF School Aviation Medicine Randolph Field, TX AF- 41-(128)-31*, 1951.
- [9] M. Hagan, *Matlab Neural Networks Tool Box*, 1993, pp. 8-11.
- [10] S. Hara, Y. Yamamoto, T. Omata and M. Nakano, "Repetitive Control Systems: A New Type of Servo System for Periodic Exogenous Signals," *IEEE Trans. on Auto. Control*, Vol. 33, No. 7, 1988, pp. 659-668.
- [11] K. J. Hunt et al, "Neural Networks for Control Systems -- A survey," *Automatica*, Vol. 26, No. 8, 1992, pp. 1085-1112.
- [12] E. Kreyszig, *Advanced Engineering Mathematics*, Second Edition, John Wiley & Sons Inc., 1968.
- [13] F. L. Lewis, K. Liu and A. Yesildirek, "Neural Net Robot Controller with Guaranteed Tracking Performance," *IEEE Trans. on Neural Networks*, Vol. 6, No. 3, 1995, pp. 703-715.
- [14] A. Isidori, *Nonlinear Control Systems*. Berlin: Springer-Verlag, 1989.
- [15] A. Karakasoglu, S. I. Sudharsanan and M. K. Sundareshan, "Identification and Decentralized Adaptive Control Using Dynamical Neural Networks with Application to Robotic Manipulators," *IEEE Trans. on Neural Networks*, Vol. 4, No. 6, 1993, pp. 919-930.
- [16] E. N. Lorenz, "Atmospheric Predictability as Revealed by Naturally Occurring Analogues," *J. of the Atmospheric Science*, Vol. 26, 1969, pp. 636-646.
- [17] B. Kosko and S. Isaka, "Fuzzy logic," *Scientific American*, Vol. 269, No. 1, 1993, pp. 76-81.
- [18] K. S. Narendra and A. M. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks," *IEEE Trans. on Neural Networks*, Vol. 1, No. 1, 1990, pp. 4-27.

- [19] W. T. Miller, R. S. Sutton and P. J. Werbos (Eds), *Neural Networks for Control*, The MIT Press (Third Printing) 1992.
- [20] A. W. Moore, "Efficient Memory-based Learning for Robot Control," Ph. D. Thesis, Tech. Rep. 229, Computer Laboratory, University of Cambridge, Oct., 1990.
- [21] M. M. Polycarpou and A. T. Vemuri, "Learning Methodology for Failure Detection and Accommodation," *IEEE Control Systems*, Vol. 15, No. 3, 1995, pp. 16-25.
- [22] Z. R. Novakovic, "The Principle of self-support in Robot Control Synthesis," *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 21, No. 1, 1991, pp. 206-220.
- [23] S. Schaal and C. G. Atkeson, "Robot Juggling: Implementation of Memory-based Learning," *IEEE Control Systems*, Vol. 14, No. 1, 1994, pp. 57-71.
- [24] H. Seraji, "Decentralized Adaptive Control of Manipulators: Theory, Simulation and Experimentation," *IEEE Trans. on Robotics and Automation*, Vol. 5, No. 2, 1989, pp. 183-201.
- [25] S. N. Singh, W. Yim and W. R. Wells, "Direct Adaptive and Neural Control of Wing-Rock Motion of Slender Delta Wings," *AIAA J. of Guidance, Control and Dynamics*, Vol. 18, No. 1, 1995, pp. 25-31.
- [26] J. E. Slotine and W. Li, *Applied Nonlinear Control*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1991.
- [27] C. Stanfill, and D. Waltz, "Toward Memory-based Reasoning," *Communications of the ACM*, Vol. 29, No. 12, 1986, pp. 1213-1228.
- [28] K. Steinbuch and U. A. Piske, "Learning Matrices and Their Applications," *IEEE Trans. on Electronic Computers*, Vol. 12, 1963, pp. 846-862.
- [29] M. Teshnehlab and K. Watanabe, "Self Tuning of Computed Torque Gains by Using Neural Networks with Flexible Structures," *IEE Proc.-Control Theory Appl.*, Vol. 141, No. 4, 1994, pp. 235-242.
- [30] D. L. Waltz, "Applications of the Connection Machine," *Computer*, Vol. 20, No. 1, 1987, pp. 85-
- [31] L. X. Wang, "Stable adaptive fuzzy control of nonlinear system," *IEEE Trans. on Fuzzy Systems*, Vol. 1, No. 1, 1993, pp. 146-155.
- [32] L. X. Wang, *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1994.
- [33] P. Werbos, "Beyond regression: New Tools for Prediction and Analysis in Behavioral Science," Ph. D. dissertation, Harvard University, Cambridge, MA, Aug. 1974.
- [34] T. Yamada and T. Yabuta, "Neural Network Controller Using Autotuning Method for Nonlinear Functions," *IEEE Trans. on Neural Networks*, Vol. 1, No. 4, July, 1992, pp. 595-601.
- [35] K. Youcef-Toumi and S. Reddy, "Analysis of Linear Time Invariant Systems with Time Delay," *J. of Dynamic Systems, Measurement, and Control*, Trans. of the ASME, Vol. 114, No. 12, 1992, pp. 544-555.

**MICROSTRUCTURAL DEVELOPMENT
DURING HOT DEFORMATION**

**Raghavan Srinivasan
Associate Professor
Mechanical and Materials Engineering Department**

**Wright State University
Dayton, Ohio 45435**

**Final Report for:
Summer Faculty Research Program
Wright Laboratory/Materials Directorate**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Wright Laboratory

September 1996

MICROSTRUCTURAL DEVELOPMENT DURING HOT DEFORMATION

Raghavan Srinivasan
Associate Professor
Mechanical and Materials Engineering Department
Wright State University

Abstract

The microstructure of a material during hot working depends upon the processing conditions, such as state of stress, strain, strain-rate, and temperature, during deformation. A new strategy for microstructural control using modern control theory and state space models for describing material behavior has been developed at the Process Design Branch (WL/MLIM). Experimental studies were conducted to validate this new strategy using hot extrusion of AISI 1030 plain carbon steel.

In another phase of research, microstructural development during hot deformation of 304L stainless steel was initiated. This alloy has wide applicability in chemical and metallurgical industry where toughness and high temperature corrosion resistance is required at high temperatures. Samples of 304L had been previously deformed by press forging, extrusion, and rolling over a wide range of temperatures (600 to 1200°C) and strain rates (0.001 to 100s^{-1}) to a strain of 0.7. Transmission electron microscopy (TEM) of the deformed samples was initiated in order to clearly identify the various microstructural processes (dynamic recovery, dynamic recrystallization, twinning, etc.) that occur during deformation.

MICROSTRUCTURAL DEVELOPMENT DURING HOT DEFORMATION

Raghavan Srinivasan

1. Introduction

The properties and performance of a component during service are influenced by the microstructural condition of the material; and the microstructure is, in turn, strongly influenced by the processing conditions during the manufacture of the component. Over the past several decades significant effort has been expended in establishing the performance-property-microstructure-processing relationships. Such effort continues as new materials are developed and existing materials are used in new applications.

In the following sections, two activities undertaken during the author's tenure at WL/MLIM are described. The first involved the experimental validation of a design methodology previously developed at WL/MLIM, and the second, transmission electron microscopy (TEM) of the deformed samples of 304L stainless steel to clearly identify the various microstructural processes (dynamic recovery, dynamic recrystallization, twinning, etc.) that occur under different deformation conditions. This work is expected to continue even after completion of the authors tenure at Wright Laboratory.

2. A New Approach to Microstructural Optimization

A number of empirical relationships between processing conditions, such as strain, strain-rate, and temperature, and the final microstructure in a hot formed component have been developed. However, a systematic use of these relationships through a formal application of modern control theory has been lacking. A new strategy

for systematically calculating near optimal process parameters for control of microstructure during hot deformation processes was developed at the Process Design Branch of the Wright Laboratory-Materials Directorate. This strategy, which involves developing state space models from available materials behavior and deformation process models, consists of two stages: the microstructural development stage and the process modeling stage. Both stages require analysis and optimization. From the first stage, trajectories of strain, strain-rate, and temperature, that will result in the previously identified "optimum" microstructure in the material, are obtained. A suitable processes simulation model, such as the slab method or the finite element method, is used in the second stage to obtain optimum process control parameters, such as die shapes, initial billet temperatures, and ram velocity, that best achieve the desired trajectories of strain, strain-rate, and temperature at specific regions of the deforming work-piece [1-4].

The following sections briefly describe the design methodology and its application to the hot extrusion of plain carbon steel with 0.3 wt. C (ANSI 1030.)

2.1. The Methodology

An empirical model for microstructural changes in 0.3 wt.% C steel, developed by Yada [5], is shown in table below:

Volume fraction recrystallized	$\chi = 1 - \exp\left(\ln(2)\left((\varepsilon - \varepsilon_c) / \varepsilon_{0.5}\right)^2\right)$
Critical strain for recrystallization	$\varepsilon_c = 4.76 \times 10^{-4} \exp(8000 / T)$
Plastic strain for 50% volume fraction recrystallization	$\varepsilon_{0.5} = 1.144 \times 10^{-5} d_0^{0.28} \dot{\varepsilon}^{0.05} \exp(6420 / T)$
Average recrystallized grain size	$d = 22600 \dot{\varepsilon}^{-0.27} \exp(-0.27Q / RT)$
Activation Energy and gas constant	$Q=267 \text{ kJ/mol}, R=8.314 \times 10^{-3} \text{ kJ/(mol-K)}$

The volume fraction recrystallized, χ , is essentially zero until a critical amount of strain, ϵ_c has been imposed. Beyond this critical strain the kinetics of recrystallization at any temperature T is determined by $\epsilon_{0.5}$, the additional strain for 50% recrystallization. The critical strain ϵ_c , strain to 50% recrystallization, $\epsilon_{0.5}$, and the average grain size after recrystallization, d , all depend upon the imposed strain rate $\dot{\epsilon}$ and temperature T . The current state or condition of the microstructure can be described in terms of several parameters or *state variables*, namely the volume fraction recrystallized, the imposed strain, and the current temperature. A *state space model*, that describes the evolution of the state of the microstructure, can be defined by the following set of equations:

$$\begin{bmatrix} \dot{\chi} \\ \dot{\epsilon} \\ \dot{T} \end{bmatrix} = \begin{bmatrix} \frac{2 \ln 2}{(\epsilon_{0.5})^2} (\epsilon - \epsilon_c)(1 - \chi) \dot{\epsilon} \\ u \\ \frac{\eta}{\rho C_p} \sigma(\epsilon, \dot{\epsilon}, T) \dot{\epsilon} \end{bmatrix}$$

where σ is the flow stress of the material, which depends upon strain, strain-rate, and temperature as follows [6]:

$$\sigma = \frac{\sinh^{-1} \left[(\dot{\epsilon} / A)^{1/n} \exp(Q / RT) \right]}{1.15 \times 10^{-5}}$$

$$\ln A = 13.92 + 9.023 / \epsilon^{0.502}$$

$$n = -0.97 + 3.787 / \epsilon^{0.368}$$

$$Q = 125 + 133.3 / \epsilon^{0.393}$$

The strain, strain-rate, and temperature imposed on the material could vary with time. These variations, or *trajectories*, influence the microstructural state both during and at the end of deformation. In the first stage of the design methodology, in order to obtain optimal trajectories, the following cost function to obtain a final strain of ϵ_f at the *end* of

deformation and a recrystallized grain size of d_f during the deformation was formulated.*

$$J = 10(\varepsilon(t_f) - \varepsilon_f)^2 + \int_0^{t_f} (d(t) - d_f)^2 dt$$

Using methods described elsewhere [7], starting with an initial guess, it was possible to identify the optimum trajectories of strain, strain-rate, and temperature for obtaining three different grain sizes of 26 μ m, 30 μ m, and 15 μ m after deformation to a final strain of 2.0. These optimum trajectories are shown in Figure 1. The initial temperature of the billet for the three cases were 1273K, 1273K, and 1223K, respectively. The temperature increases with time due to deformation heating. In order to maintain grain size a constant, the optimization routine increases the strain rate simultaneously. The initial grain size of the billet was assumed to be 180 μ m. In each case, the grain size decreases to the design grain size upon crossing the critical strain of about 0.25. Subsequently, the optimization routine ensures the grain size remains constant.

The trajectories identified above are independent of the process, i.e., they are valid for any method of deforming the material, whether it be forging, rolling, extrusion, or any other process. The second stage of the design methodology involves determining specific parameters related to a given process that will impose these optimal trajectories on the entire deforming body or at certain critical regions. In the case of extrusion, these parameters are the die profile and the velocity of the ram. For extrusion, it is also possible to analytically determine the die profile and ram velocity for

* Yada's equations indicate that recrystallization occurs very rapidly, with very little strain (less than about 0.01 at temperatures in the range of 1250K) required beyond critical strain to cause complete recrystallization. Due to this "step-function" like behavior, no attempt was made to control the volume fraction recrystallized to less than 100%.

a given die entrance radius r_0 (also equal to the billet radius) and die length L , if the strain trajectory $\varepsilon(t)$ is known [8].

$$V_{ram} = \frac{L}{\int_0^{t_f} \exp(\varepsilon(t)) dt}$$

where t_f is the duration of the extrusion process. The die shape is described by pairs of points (r,y) which give the radius r of the die at an axial distance y from the die entrance.

$$r = r_0 \exp(-\varepsilon(t) / 2); \quad y = V_{ram} \int_0^t \exp(\varepsilon(t)) dt$$

Figure 2 gives the optimal die profile for achieving final grain sizes of 26 μ m, 30 μ m, and 15 μ m using this approach. The corresponding optimal ram velocities were 8.43mm/s, 5mm/s, and 25.1mm/s. Note that the die profiles are almost identical for all three cases. This means that the same die can be used to achieve different recrystallized grain sizes simply by changing the ram velocity and the initial billet temperature.

2.2. Experimental Validation

The optimization methodology was validated by means of extrusion experiments on a 6000 kN Lombard horizontal extrusion press at Wright Patterson AFB. Two extrusion experiments were conducted to yield grain sizes of 26 μ m and 15 μ m. A die of the shape shown in Figure 2 was fabricated out of H-13 tool steel. Billets of AISI 1030 steel were heated to the recommended temperatures, and extruded at the ram speeds mentioned above. For these experiments the extrusion press chamber, die, and billet follower block had been preheated to 533K before the start of the extrusion. The

extrudates were allowed to drop directly into a water quench tank upon emerging from the die, with a delay of less than 5 seconds.

Specimens were cut out from the lead and tail end of the extrudates, sectioned along a diametral plane, metallographically polished and etched to reveal the prior austenite grain boundaries. Figures 3 and 4 show typical microstructures along the centerline of the extrudate. Grain size measurements on the two extrudates show that the grain size is approximately $27\mu\text{m}$ and $16\mu\text{m}$, compared to the design objective of $26\mu\text{m}$ and $15\mu\text{m}$, respectively.

3. TEM of 304L Stainless Steel

Austenitic stainless steels are widely used in chemical and metallurgical processing industry due to their high toughness and corrosion resistance at elevated temperatures. These alloys contain chromium and nickel as the primary alloying elements. A typical example of such an alloy is 304L with a composition (wt%) of 18-20 % Cr, and 8-10.5% Ni. The L indicates low interstitial alloying contents and requires a maximum of 0.08% C. The chromium is required for developing a corrosion resistant oxide layer, and the nickel ensures the FCC gamma phase (austenite) is retained at room temperature. The limitation on carbon content is to prevent the depletion of chromium by precipitation of carbide particles along grain boundaries. Since the FCC structure is retained to room temperature, austenitic stainless steels can be cold, warm and hot formed.

Samples of 304L stainless steel had been previously deformed under different processing conditions [9]. An analysis of the flow curves and optical microscopy of the samples have tentatively identified the microstructural mechanisms that operate at

different temperatures and strain rates. However, a definite identification of regimes where specific processes, such as dynamic recrystallization (DRx), dynamic recovery (DRy), deformation twinning, martensite formation, and ferrite formation, occur, had not been completed. From the matrix of test conditions, samples deformed at a strain rate of 0.1s^{-1} at temperatures of 700, 800, 900, 1000, 1100, and 1200°C , and those deformed at 1100°C , at strain rates of 0.001, 0.01, 0.1, 1, 10, and 100s^{-1} were selected for TEM examination. Initial results indicate that the microstructure changes from deformation twinning to dynamic recovery to dynamic recrystallization, with increasing temperature of deformation. Additional work is required to complete this characterization.

4. Summary

The author was involved in two activities during the course of his tenure at the Process Design Branch (WL/MLIM) of the U.S. Air Force Wright Laboratory - Materials Directorate: (i) experimental validation of a new design methodology for microstructural optimization, and, (ii) transmission electron microscopy of deformed samples of 304L stainless steel.

The first activity resulted in several publications, listed as References 1-4, as well as the following disclosure and record of invention:

"Optimization and control of microstructure development during hot metal working: a new technique using modern control theory," by J.C. Malas, W.G. Frazier, E.A. Medina, V. Seetharaman, S. Venugopal, R.D. Irwin, W.M. Mullins, S.C. Medeiros, A. Chaudhary, R. Srinivasan, August 1996.

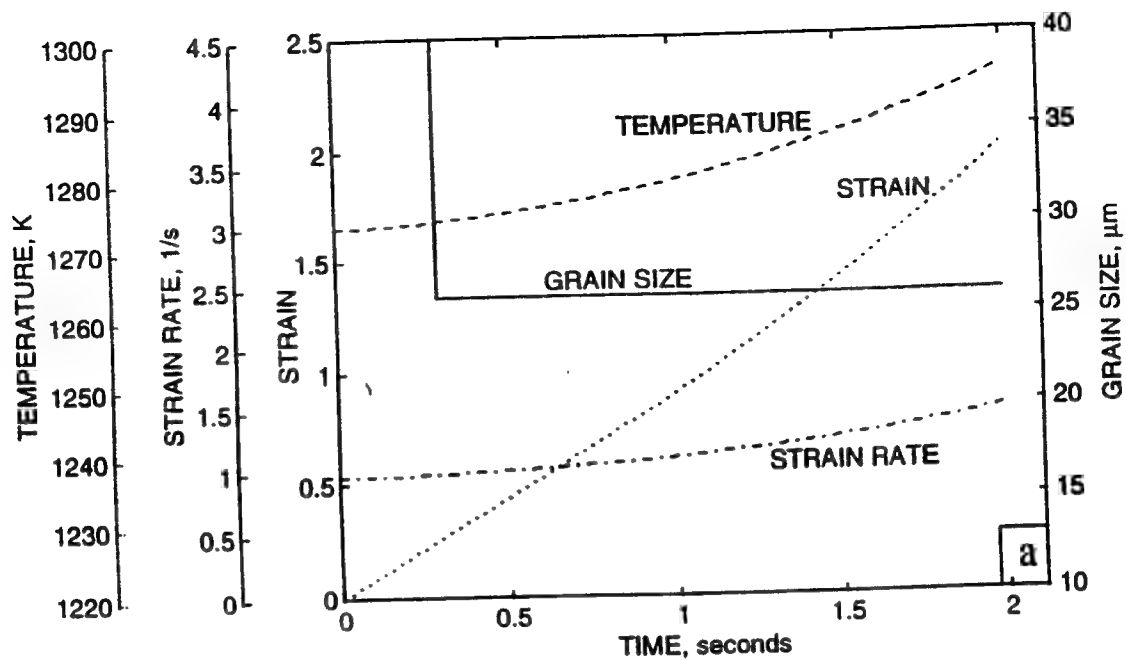
The second activity, namely transmission electron microscopy of 304L stainless steel, was initiated. This work is expected to continue beyond the author's tenure at Wright Laboratory.

Acknowledgments

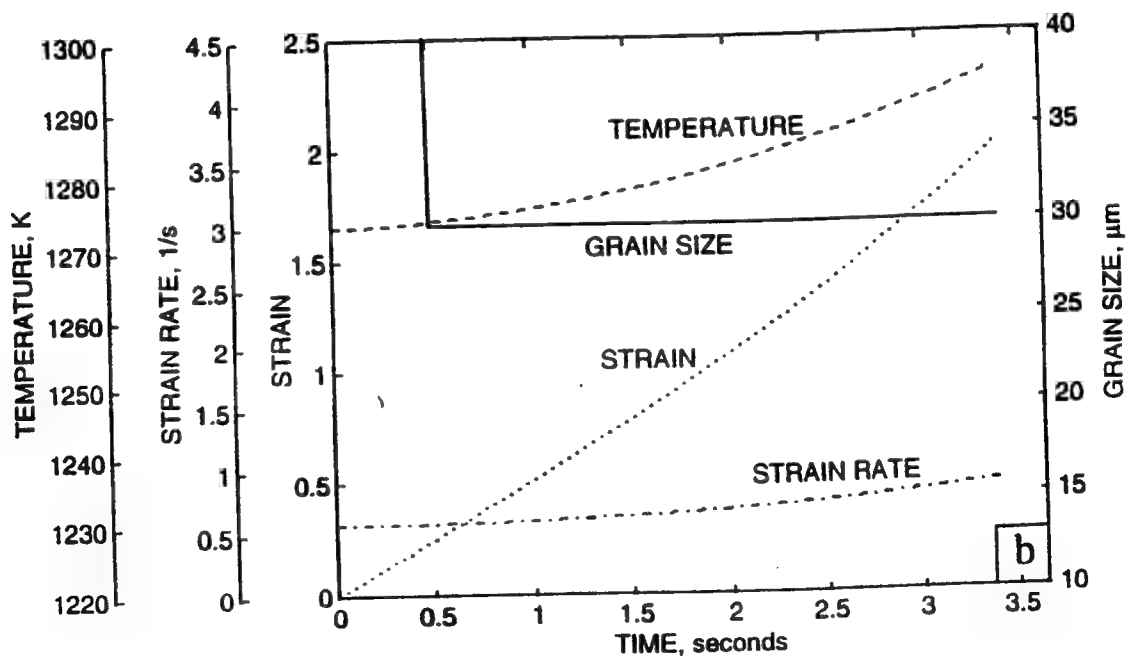
The author wishes to thank Drs. Jim Malas, S. Venugopal, Bill Mullins, Anil Chaudhary and Garth Frazier, and Messrs. Steve Medeiros and Enrique Medina for the many stimulating discussions which have resulted in a very interesting and productive Summer spent at the Wright Laboratories.

References

1. J.C. Malas, A. Chaudhary, W.M. Mullins, E.A. Medina, S. Venugopal, S. Medeiros, R.D. Irwin, W.G. Frazier, and R. Srinivasan, "Optimization of Microstructure Development: Application to Hot Metal Extrusion," PD Vol. 75, *ESDA Proceedings*, Vol. 3, pp. 125-135, ASME New York, 1996.
2. S. Venugopal, E.A. Medina, J.C. Malas, S. Medeiros, W.G. Frazier, W.M. Mullins, and R. Srinivasan, "Optimization of Microstructure Development during Deformation Processing using Control Theory Principles," accepted by *Scripta Materialia*, September 1996
3. J.C. Malas, W.G. Frazier, S. Venugopal, E.A. Medina, S. Medeiros, R. Srinivasan, R.D. Irwin, W.M. Mullins, and A. Chaudhary, "Optimization of Microstructure Development during Deformation Processing using Control Theory," submitted to *Metals and Materials Transactions*, August 1996.
4. E.A. Medina, S. Venugopal, W.G. Frazier, S. Medeiros, W.M. Mullins, A. Chaudhary, R.D. Irwin, R. Srinivasan, and J.C. Malas, "Optimization of Microstructure Development: Application to Hot Metal Extrusion," submitted to *J. Materials Engineering and Performance*, August 1996.
5. H. Yada, *Proc. Int. Symp. Accelerated Cooling of Rolled Steels*, Conf. of Metallurgists, CIM, Winnipeg, MB, Canada, Aug. 24-26, G.E. Ruddle and A.F. Crawley, Eds., Pergamon Press, Canada, pp. 105-20, 1987.
6. A. Kumar, K.P. Rao, E.B. Hawbolt, and I.V. Samarasekera, The Application of Constitutive Equations for Use in the Finite Element Analysis of Hot Rolling Steel, *Unpublished Research*, 1987.
7. W.G. Frazier, "Robust Control Techniques for Hot Deformation Processes," Contributive Research and Development, Volume 228, SYSTRAN Corp. Final Report, Task 178, Contract F33615-90-C-5944, 1995.
8. R. Srinivasan, J.S. Gunasekera, H.L. Gegel, and S.M. Doraivelu, Extrusion Through Controlled Strain Rate Dies, *J. Material Shaping Technology*, Vol. 8, No. 2, pp. 133-141, 1990.
9. S. Venugopal "Deformation Processing of Austenitic Stainless Steels," Ph.D. Thesis, University of Madras, 1993.



(a) 26 μm grain size



(b) 30 μm grain size

Figure 1: Optimum trajectories of process parameters to obtain a final grain size of (a) 26 μm , (b) 30 μm , and (c) 15 μm in AISI 1030 steel.

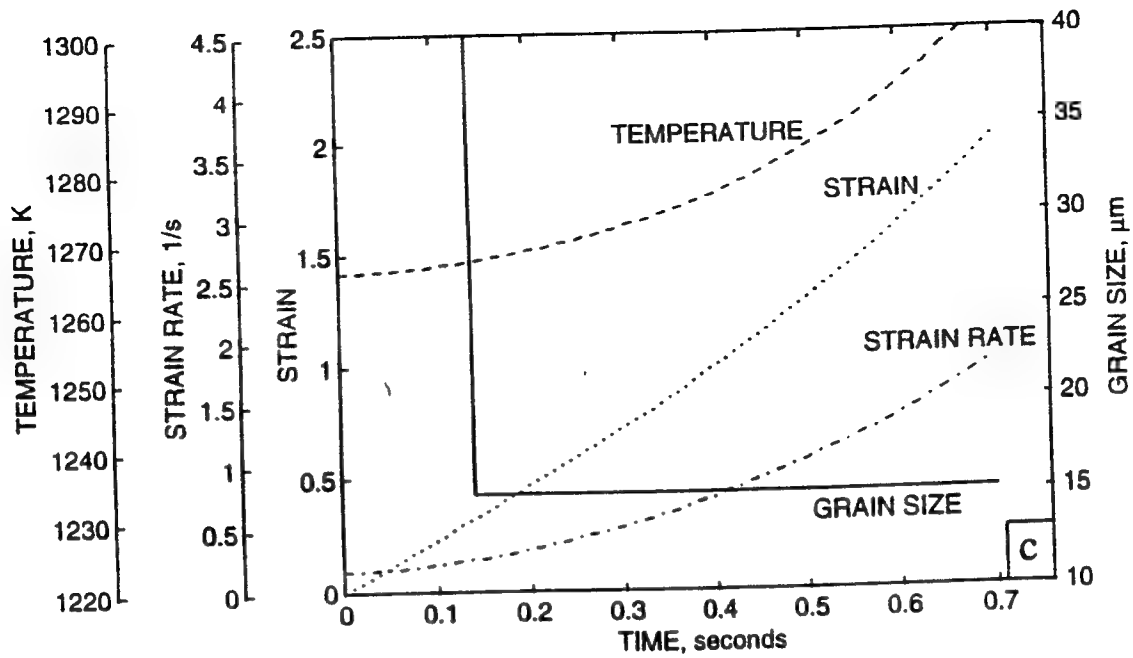


Figure 1(c) 15μm grain size

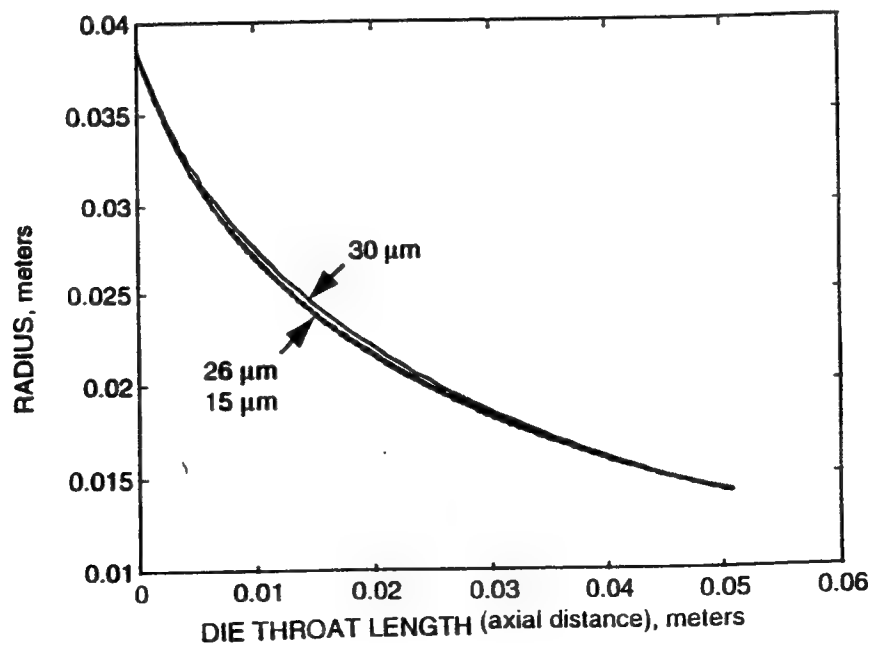
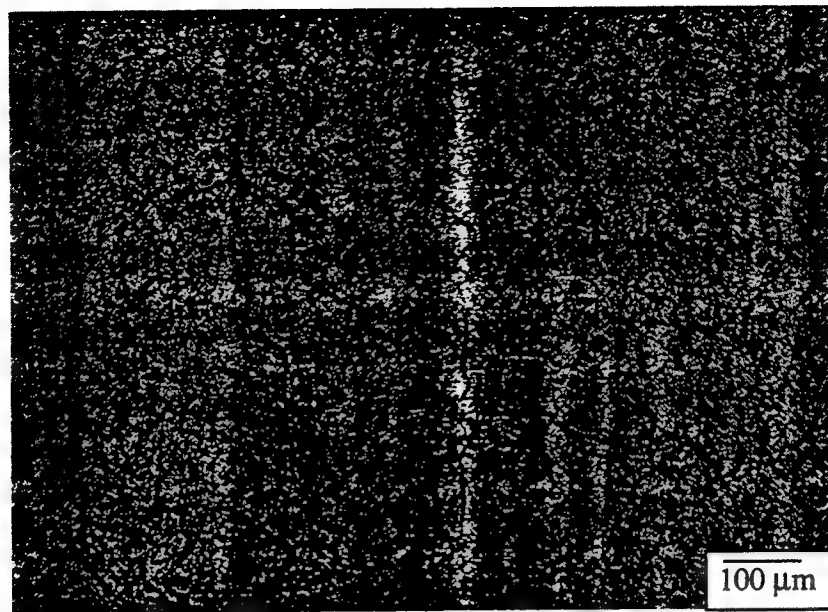
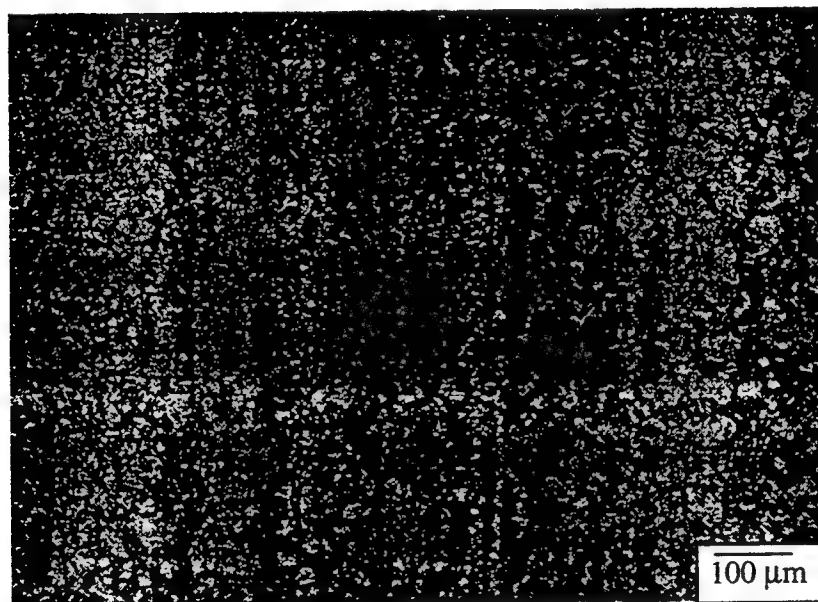


Figure 2: Die profiles for obtaining trajectories shown in Figure 1.

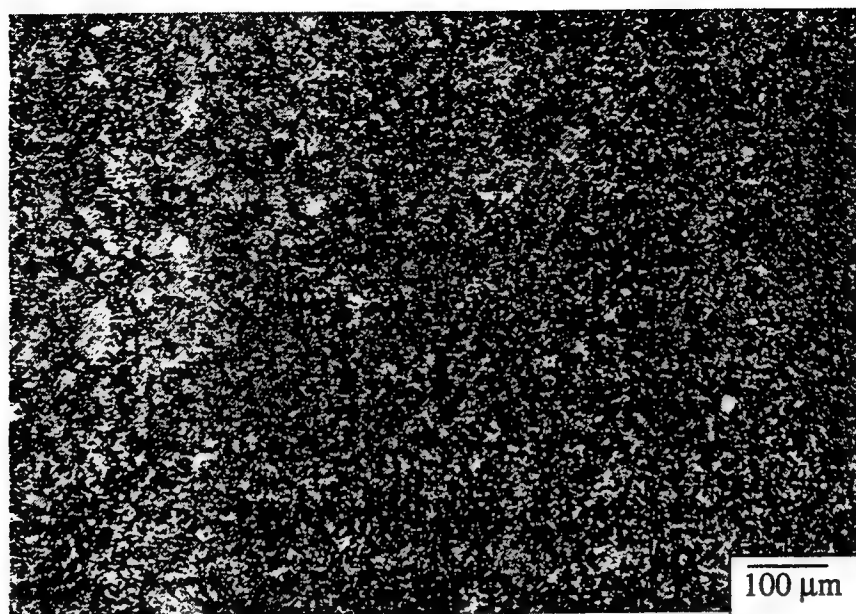


(a) Lead End

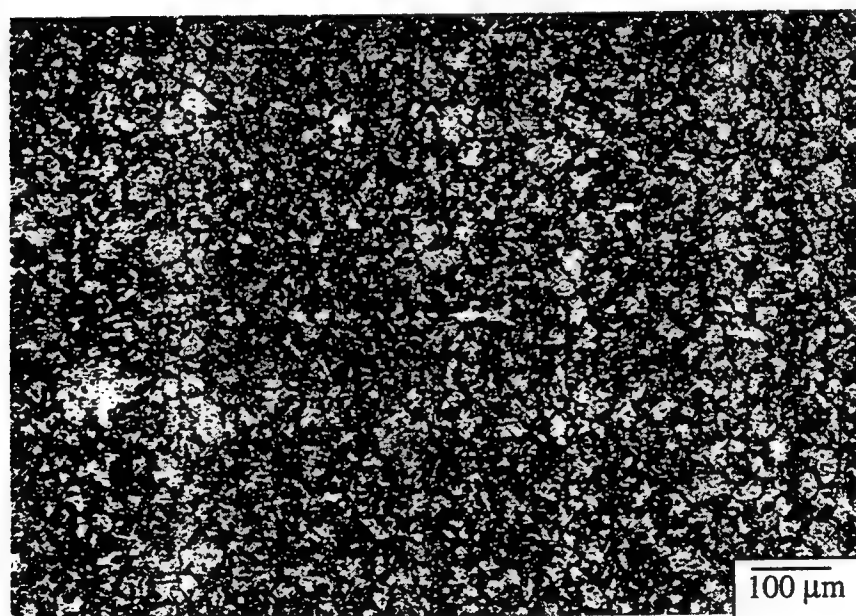


(b) Tail end

Figure 3: Typical microstructure from AISI 1030 steel extruded to obtain a final grain size of 26 μm.



(a) Lead End



(b) Tail end

Figure 4: Typical microstructure from AISI 1030 steel extruded to obtain a final grain size of 15 μm.

FEATURE SELECTION FOR ATR
NEURAL NETWORK APPROACH

Janusz A. Starzyk
Professor

School of Electrical Engineering
and Computer Science

Ohio University
Athens, OH 45701

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Laboratory

August 1996

FEATURE SELECTION FOR ATR NEURAL NETWORK APPROACH

Dr. Janusz A. Starzyk
School EECS, Ohio University, Athens, OH 45701

Abstract:

This report investigates independent feature selection in neural networks for high resolution radar target recognition. Radial basis functions and wavelet transforms are used to preprocess the radar signal data. Data preprocessing is used to minimize the effect of noise, phase shift, and scale change of the radar signal. A class of nonorthogonal classifiers is defined and their classification properties are investigated. It is demonstrated that nonorthogonal classifiers perform better than the orthogonal ones. Then, feature selection using the mutual information is investigated. Independence of features based on the information content is defined and used to select features for synthesis of ontogenic neural networks. Simulation results using synthetically generated aircraft images showed promise for automatic target recognition.

FEATURE SELECTION FOR ATR NEURAL NETWORK APPROACH

Janusz A. Starzyk

Introduction

Automatic target recognition requires an accurate recognition algorithm which can be implemented on real time hardware. This problem can be efficiently solved on a dedicated neural network (NN) trained off line with an extensive data base to perform a pattern classification task on the incoming signals. Neural networks use many paradigms and data organizations which are potentially useful for target recognition.

Radial basis functions (RBF) were shown to provide an arbitrary close approximation to any continuous function [Girosi]. In addition [Chen] demonstrated that any continuous function can be used as an activation function in RBF neural networks (RBFN), as long as it is not an even polynomial. Recently [Zhao] used RBFN for automatic target recognition based on Ku-band radar range profiles of three military aircraft. They demonstrated that by using Fourier transform and a non-coherent amplitude averaging, stable and shift invariant features were obtained which significantly improved the classification quality. In addition, classification based on RBFN was more accurate than using minimum-error Bayesian classifiers.

In this work independent features were selected to construct minimal classifiers for ATR. The input transformation space includes the original signal, its amplitude range, average value, standard deviation, and additive information cost functions (Shannon entropy, log energy, and P norms.). In addition Fourier transform and Haar wavelet transform of the original signal were used to enhance the input transformation space. In this space a selection of features for nonorthogonal classifiers were performed. The classification results were compared with classification based on the orthogonal RBF classifiers obtained in the same input transformation space. NN training and test were performed using model based synthetically generated aircraft images of a high range resolution radar.

Classifiers

A fundamental problem in pattern classification is to determine class membership with the maximum statistical confidence of the correct classification decision. While this classification task can be performed with 100% probability for the trained data, there is no proof that a neural network capable of 100% recognition of the trained data yields a similarly high recognition rate for test data which were not used during the training. To the contrary, such networks are quite often unable to classify new test data and are known as over trained networks unable to generalize. In addition, the over trained neural networks are usually designed without regard to the hardware cost and result in an excessive architecture for a given task. Therefore, selecting a neural network for pattern classification and its training algorithm, one must consider tradeoffs between the request to separate members of different classes and the request to unite members of the same class, as those requests translate into contradicting hardware requirements. Unification of a large number of training data into a single decision making operation usually translates into simplification of the classification task and produces a NN capable of a useful generalization, while separation makes it more complex and reduces its generalization ability.

In order to facilitate the classification task, distinguishing target features are used. These features are selected from a predefined set of the transformation functions of the input data. Depending on the task and imagery data the separation boundaries between different classes will be crisp or fuzzy, leading to different types of neural network processing components and different types of membership functions of the resulting classification. To construct a feature selection mechanism we define a measure of performance relative to the classification task. We will represent a given pattern as an ordered set of values. For instance, a binary image can be represented by a vector of 0 and 1 values arranged in the specific scanning order or a Fourier spectrum of a radar signal can be represented by a vector of complex values at discrete frequency points.

To compare and classify patterns using such a vector representation and to preserve all the useful information before a feature selection criteria are established, we assume that in a given

classification task all vectors have the same dimensions. If needed, vector sizes may be extended to match the longest vector in the input pattern set. Vector padding, sampling, time, and frequency scaling may be used to satisfy this requirement. A linear vector space which contains all the vectors obtained from the input pattern set is called the *input space*. Input space may either contain raw input data or some functions of these data. By allowing initial transformation of the input data we have greater flexibility to formulate input space. A direct representation of the input signals and its transforms yields vectors of high dimensionality. To facilitate the classification task, the transformation of the input space to a space of a lower dimension is used and is optimized for a specific classification task.

A specific transformation of the input is referred to as the *input feature*. Using an unlimited number of features one can achieve a linear separability of any input data, which is critical for any classification problem. But this selection may lead to a costly classification procedure and hardware. In addition, it will not be useful for classification of a new data. By proper selection of the input features we can obtain better classifiers, resistant to noise and local image distortions, and applicable to different types of sensor data.

A mathematical task of checking the independence of functions or vectors is easily accomplished by using linear algebra methods. On the other hand, there is no mathematically formulated independence of features selected for the classification task, which makes design of optimum software and hardware methods for pattern recognition hard, and also makes any comparison between different classifiers difficult. In this research we define a way to check the independence of features used for pattern classification. By using independent features, a systematic way of synthesizing minimum classifiers will be derived, which will permit the design of optimized architectures of the resulting NN hardware.

Since the individual features will be defined to optimize the classification task, the obtained classifiers will be optimum in the sense of the selected transformations, so they are expected to outperform any ad-hoc classification technique. The resulting optimized classifiers will perform

their classification task with a high tolerance to deviations within a specific class. They will be able to avoid incorrect classifications and minimize the false alarms better than ad-hoc designs.

Feature selection for NN pattern recognition

:

To learn a nonlinear mapping from the input space to the output space one needs to consider independent transformations of the input space. Such transformations can be easily obtained using either a complete set of orthogonal functions, in which orthogonality guarantees transformation independence, or using a successive approximation of the learned mapping, in which each successive transformation is found by orthogonalizing the error of the existing fit. Similar to this operation is orthogonalization of vectors in a linear vector space. By using Gram-Schmidt orthogonalization, we successively remove the components of the input vectors which are linearly dependent on the selected orthogonal set of vectors. In order to obtain an efficient classifier a similar operation needs to be defined on a set of input transformations (features). This operation would tell us how independent a classification based on a selected feature is as compared to the previously selected features. By introducing this operation and a measure of separation ability we will define a tool for selection of the dominant features for building the optimum classifiers.

To this end, let us define a *feature* f mathematically as an ordered pair (F, Ω) of a nonlinear transformation F and a proper subset of its output space. We define a *feature domain* D as a subset of the domain of the transformation F which is mapped into Ω , and a *feature sample set* S as a subset of the input training data included in D . For instance a cluster of points in the input space is defined by a feature in which a nonlinear transformation F is the Euclidean distance from the cluster center and the proper subset Ω is the closed interval $[0, R]$, where R is the cluster radius. The feature domain D is a sphere with radius R located at the cluster center, and the feature sample set S is equal to the set of input points included in D .

Notice, that for a given transformation F we can define infinitely many features by simply modifying the subset Ω . An entire classification task can be based on a single transformation F paired with different sets Ω . For example a clustering method can be used to classify a number of patterns based on the nearest neighbor rule and a number of clusters defined in the input space.

Definition: A feature f_m is **covered** by the features f_1, f_2, \dots, f_k iff $D_m \subset D_1 \cup D_2 \cup \dots \cup D_k$.

Definition: A set of features $\Phi = \{f_1, f_2, \dots, f_n\}$ is **independent** if none of its elements can be covered by others.

Consider a set of *training data* T used for the task of pattern recognition. This set is composed of subsets of vectors from different classes $T = C_1 \cup C_2 \cup \dots \cup C_c$, where C_i is a set of vectors from the class I . For simplicity we will use the same symbol to represent a class and its set of input vectors. Without loss of generality let us assume that all the classes are disjoint, i.e. $C_i \cap C_j = \emptyset$ for $i \neq j$.

Definition: A feature $f(C_i)$ is a **distinguishing feature** of class C_i if its domain includes only the input vectors from the class C_i .

Notice, that whether a feature is a distinguishing one or not depends on the complete training set T . Adding new training data may change a distinguishing feature to become nondistinguishing and removal of a training data may change a nondistinguishing feature to a distinguishing one.

Definition: A set of distinguishing features $\Phi(C_i) = \{f_1(C_i), f_2(C_i), \dots, f_n(C_i)\}$ is an **orthogonal classifier** for class C_i if the sum of domains of its features includes the set C_i

$$C_i \subset D_1 \cup D_2 \cup \dots \cup D_n$$

Class C_i becomes the domain of its orthogonal classifier. The classifier is a *minimal orthogonal classifier* if the distinguishing features are independent.

Features can be treated as logic functions (crisp or fuzzy) defined on the entire input space. The logic value of such a function $f(x)$ at a given point x is true if $x \in D$. In the orthogonal classifier we can identify the classification process with evaluation of the logic expression

$$\phi(x) = f_1(x) + f_2(x) + \dots + f_n(x)$$

with $x \in C_i$ iff $\phi(x)=1$, where f_1, f_2, \dots, f_n are all the features of the classifier ϕ .

In general, to classify input data we can use a set of orthogonal classifiers defined for all classes. Using orthogonal classifiers will yield 100% recognition for the training data. This is the case in a clustering technique when nonoverlapping clusters are used to classify the input data. In addition, orthogonal classifiers will result in a simple two level structure of the classification neural network. However, the recognition rate for new data may be significantly smaller, as quite often the orthogonal classifiers are based on a large number of independent features, which leads to small feature domains and poor generalization ability. In addition, as a result of a large number of features used, the hardware requirements for the neural network based on orthogonal classifiers are high. The network will use features defined in the input space or the transformed input space, and it will have a two level structure with a high demand on wiring the input data. Orthogonal classifiers can be designed independently from each other since their features sample sets are disjoint. In addition, the classification decision is made based on a single feature being true. We could represent the classifier's decision as being based on the output of an OR gate which takes all its feature values as inputs. Combinational classifiers will make their decision based on combination of features (or their complements) being true, so they will be designed using more complex logic structures of decision making process. Probabilistic or fuzzy classifiers can also be designed by using fuzzy logic instead of the binary one.

Nonorthogonal classifiers, contrary to orthogonal classifiers, do not have to have disjoint domains. They can be constructed using different features and some of them may have better classification properties than orthogonal classifiers. This point will be demonstrated using so called sequential classifiers.

:

Definition: A *dominating distinguishing feature* is a distinguishing feature with the highest ratio of the cardinality of its feature set over the cardinality of the associated class set.

In order to design a simple classifier we must find a set of independent features of minimum cardinality which distinguishes all classes. In what follows, feature selection for a *sequential classifier* is described. It will be demonstrated that this classifier can produce better results than an equivalent orthogonal classifier.

Sequential feature selection (SFS) is a process described as follows:

1. Select a dominating feature f_i
2. Remove feature's sample set S_i from the input training set T and modify T , $T=T-S_i$.
3. If the training data set T is not empty go back to 1., otherwise stop.

As a result of the SFS process a sequence of dominating features (f_1, f_2, \dots, f_d) associated with various classes are obtained. By the process organization, all the dominating features obtained are independent. Notice that since SFS is performed with modification of the input training data, the obtained dominating features are no longer distinguishing features of their corresponding classes for the original input training data. However, distinguishing features can be easily generated from the obtained sequence by logic operations using the preceding features in a sequence.

Lemma: If the feature f_m obtained in the SFS process, was selected as a dominating feature of the class C_i , then $\xi_m = f_m \prod_{k=1}^{m-1} \bar{f}_k$ is a distinguishing feature of the class C_i in the input space with the original training data.

The equation defining ξ_m is expressed by a logic function of dominating features. In fact, the expression for ξ_m a distinguishing feature of class C_i , can be further simplified by removing all references to f_k which are dominating features of the same class and by removing all references to f_k for which $S_k \cap S_m = \emptyset$ for the original training data.

By logically adding all the distinguishing features of class C_i we define its classifier $\Phi(C_i) = \{\xi_1(C_i), \xi_2(C_i), \dots, \xi_n(C_i)\}$, and classify $x \in C_i$ iff $\Phi(x)=1$, using the logic expression

$$\phi(x) = \xi_1(x) + \xi_2(x) + \dots + \xi_n(x)$$

where $\xi_1, \xi_2, \dots, \xi_n$ are all the features of the classifier Φ .

A sequential classifier is an example of a nonorthogonal classifier. In general, it has better generalization ability than an orthogonal classifier. Sequential classifiers are just one example of nonorthogonal classifiers. They result in a multilevel NN structure in which both number of neurons, processing layers and the overall organization is a function of the input data, which is a characteristic feature of the ontogenic NN [Ensley].

Since an input vector may or may not exhibit an individual feature, we can design combinational classifiers in which a decision regarding an input classification is expressed by a combinational logic function which depends on several features. This leads to the construction of pattern recognition neural networks in which classification decisions are made by a network of logic gates. Such classifiers will be extremely hardware efficient, as a multilevel logic synthesis could be used to synthesize its structure for optimum cost and performance. Combinational classifiers will be

designed using complex logic structures of the decision making process. This will yield a multilevel NN structure. Probabilistic or fuzzy classifiers can also be designed by using fuzzy logic instead of binary logic. Design of combinational classifiers will be based on feature selection using the mutual information measure.

:

Mutual information measure

A sequential classifier is a special case of a more general combinational classifier. In a combinational classifier, independent features can be selected in a number of ways. It is our aim to select these features in a way that yields optimum neural networks structures. A natural way of achieving this objective is to select a classifier which contains a minimum number of features satisfying a specified selection criteria. For instance, many orthogonal classifiers (Parzen window estimator, Karhunen-Loeve transform, scatter matrix, Bhattacharyya distance) are minimal in the sense of the specified selection criteria. In this section, feature selection based on the mutual information measure is investigated. It yields a "maximally informative" set of features which minimizes the initial uncertainty in the object class.

It is expected that this approach yields the minimum number of features to perform the classification task and, as a result, minimal NN classifier hardware will be generated. The mutual information between a feature f and a set of classes G is defined as follows [Battiti]:

$$I(f, C) = \sum_{C \in G} P(C, f) \log \frac{P(C, f)}{P(C) P(f)}$$

where $P(C, f)$ stands for the joint probability for class C and feature f , $P(C)$ is the class probability, and $P(f)$ is the feature probability. In general, the conditional entropy, which measures the uncertainty in the object class, will be reduced when a new feature is added to the classifier. It remains unchanged if and only if $P(C, f) = P(C) P(f)$, in which case the feature does not bring new information and should not be used in the classifier. [Battiti] presented convincing arguments why

the mutual information measure is useful in feature selection for object classification. His selection procedure uses a sequence of features f_i which maximizes $I(f_i, C)$. He realized that this selection may introduce features which are strongly dependent and in spite of having large $I(f_i, C)$ values their contribution to the classification problem may be much less than expected. To alleviate this problem he is using feature selection that maximizes

$$I(f, C) - \beta \sum_{f_p \in F_p} I(f, f_p)$$

where F_p is the set of the previously selected features, $I(f, f_p)$ is the mutual information (which measures dependence) between a candidate feature f and the already selected features f_p , and β is a parameter between 0 and 1 which regulates the relative importance of the mutual information between features f and f_p .

The problem with this approach is that β is arbitrarily selected and cannot correctly remove the amount of mutual information between features. The major reason why no single value of β can be found is that all previously selected features may be mutually dependent. The more dependence there is between previously selected features, the smaller the β value must be used. But this, in turn, is very much case dependent. What is more important, features selected using this criteria may be completely dependent on features previously selected and will not contribute to information increase. As a result the obtained classifier is not minimal from an information theory point of view. In addition, the method gives no clue as to how many features should be selected to reach the optimum level of information accuracy possible with a given set of features.

Motivated by the above deficiencies of the mutual information based feature selection a new method called the maximum information increase feature selection (MIIFS) was developed and tested on a selected set of input features. The method is computationally efficient and provides an optimum feature selection based on the exact information measure. In this method, both feature domain and its complement are considered in reaching the classification decision. As a consequence, each feature partitions the input space into two subspaces. If several features are considered the input space is partitioned into a number of subspaces. Each subspace is included in a unique combination of feature

domains or their complements. In the MIIFS approach, the mutual information is computed based on these subspaces. The mutual information between a set of features Φ and the set of classes G is computed from

$$I(\Phi, G) = \sum_{s \in S} \sum_{C \in G} P(C, s) \log \frac{P(C, s)}{P(C) P(s)}$$

where S is the orthogonal sum of all subspaces created by intersection of feature domains and their complements. First, the mutual information is computed for each feature with $S = D \oplus \bar{D}$, and $\Phi = f$. A feature f with the largest value of information $I(\Phi, G)$ is selected. The input space is divided into two subspaces. Next, the feature space is searched for a new feature which maximizes the information increase

$$\Delta I = I(\Phi_n, G) - I(\Phi_{n-1}, G)$$

where $I(\Phi_n, G)$ is the mutual information between new set of features Φ_n and the set of classes G . Φ_{n-1} stands for the previously selected set of features. $I(\Phi_n, G)$ is computed on a new set of subspaces which are created by intersecting a new feature domain and its complement with the previously obtained set of subspaces. Notice, that if Φ_n has n features, then up to 2^n subspaces are created. Feature selection continues until the maximum value of ΔI is less than a specified threshold. This method produces a minimum set of independent features optimized from the information theory point of view. If the information threshold is set to zero, the method produces set of features capable of 100% recognition of the trained data. Since the method minimizes the number of features selected, it is also capable of good generalization. Although the method can be used to achieve 100% recognition, the statistical confidence in the classification results may be lowered by a small increments of information added by features selected later in the process. This is a direct result of the error in estimating the mutual information I which differ from the true value of the information \bar{I} represented by the set of features by ΔI , where

$$\Delta I = I - \bar{I} \approx \frac{1}{2N} (K_c K_s - K_c - K_s)$$

and N is the total number of training samples, K_c is the number of classes, K_s is the number of subspaces. An additional issue, that must be investigated to determine the confidence in the classification result, is to determine the statistical likelihood that a new sample may be of a different class than samples represented by the feature domain (subspace) which correspond to the given combination of features. This is directly related to the number of training data in the given subspace.

Classification using slicing approach

In this work sequential classifiers are used to demonstrate a potential utilization of nonorthogonal classifiers for automatic target recognition. A simple feature selection method was used to demonstrate that the nonorthogonal classifiers work even with the simplest thresholding features. It is expected that when used with a more elaborate feature selection process nonorthogonal classifiers will maintain their advantage over the orthogonal classifiers. The experiments described in this paper used synthetically generated high range resolution radar (HRR) data. The full data set has six air-to-air targets. Each target set consists of 1071 range profiles with each range profile consisting of 128 range bins. The target pose is head-on with a range of azimuth of $\pm 25^\circ$ and elevations of -20° to 0° . From the complete data set a training and test set were extracted. Each of the training and test sets consisted of 60 randomly selected range profiles for each target. For each target in these sets the same azimuth and elevation range profiles were used. This yielded 360 range profiles for each data set. The range profiles were selected such that the training and test sets were disjoint. The reason that these data sets are so small is to speed the computations for a proof of concept. In actual usage, much larger training and test sets would be used. It should also be noted that it is very possible that the test data ranges outside the training data set. Thus causing the methodology to extrapolate and may result in lower recognition rates.

The input space includes the original signal, its amplitude range, average value, standard deviation, and additive information cost functions (Shannon entropy, log energy, and P norms.). In addition, a Fourier transform and a Haar wavelet transform of the original signal were used to enhance the input space. In this space, a selection of features for orthogonal and nonorthogonal classifiers were performed. Fig. 1(a) shows an example of a raw signal data, and Fig. 1(b) its wavelet transform.

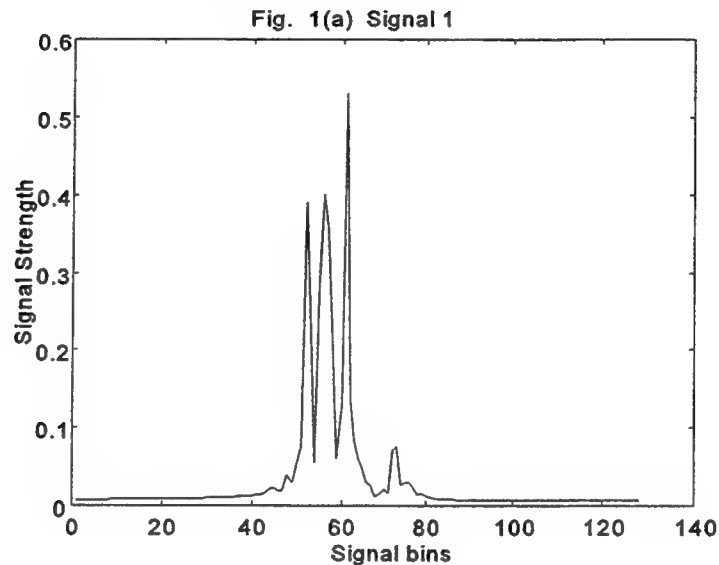


Fig. 1(b) Signal 1 and its Haar Wavelets

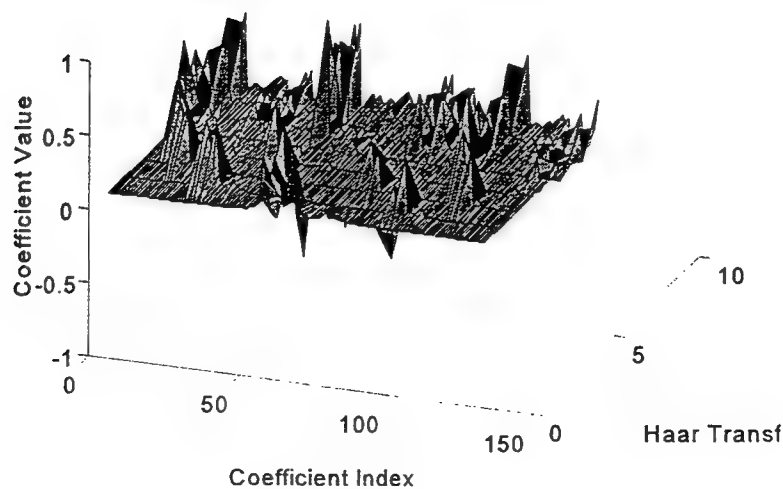


Fig. 1 (a) A sample HRR signal, (b) Haar wavelet of the selected signal.

Feature selection was based on the *slicing approach* in which input data were projected onto one dimensional subspaces. In these subspaces intervals which include input vectors from different classes were found. These intervals define slices in the multidimensional input space which contain only the samples of a single class. Features were selected based on the cardinality of input vectors included in these slices. Fig. 2 illustrates maximum cluster sizes for different transformation functions of the input data. We may observe that the best features are based on the signal transforms rather than on the raw data (represented by features 7-134.) This is in agreement with other research results in ATR, which indicate that preprocessing may enhance the classifiers' recognition ability.

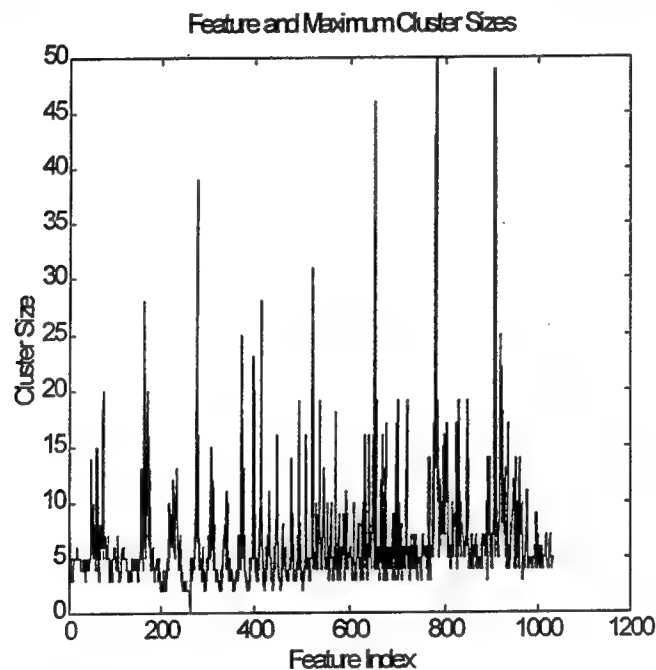


Fig.2 Different features and their maximum cluster sizes.

An orthogonal classifier was obtained by using an equal number of distinguishing features for each class. Distinguishing features were selected based on the largest sample sets. A sequential classifier was used as an example of nonorthogonal classifier. Distinguishing features were also based on the largest sample sets. Since sequential classification procedure changes the input space, after each step

the projection intervals and cardinality of input vectors in various slices changed, effecting the choice of distinguishing features. In general, the obtained sequence of features may have a nonmonotonic cardinality.

After feature selection was completed a RBFN was obtained for each classifier. This results in three layer NN for both orthogonal and nonorthogonal classifier. This was done to demonstrate differences in the classification performance of the two classifiers which result from differences in the feature selection process, rather than NN organization. Normally, nonorthogonal classifiers will be realized in a multi layer NN which may lead to simpler hardware. The results of classification are shown in Table 1.

Target number	Training Data				Test Data			
	Orthogonal		Sequential		Orthogonal		Sequential	
	recogn. rate %	error rate %	recogn. rate %	error rate %	recogn. rate %	error rate %	recogn. rate %	error rate %
1	87	0	83	0	83	0	85	0
2	53	0	60	3	28	18	32	8
3	50	0	63	0	40	2	73	3
4	37	0	67	3	25	18	53	15
5	28	0	78	2	12	0	48	10
6	28	0	48	2	18	7	35	15

Table 1. Classification results based on orthogonal and sequential classifiers

Notice, that the classification results were obtained with a small number of simple feature functions. The sequential classifier used 29 features and the orthogonal classifier used 30 features to classify targets into 6 different classes. Typically, a much larger selection of features or much more complex features are used. In addition, the training set was deliberately small to provide a proof of concept for the enhanced performance of nonorthogonal classifiers.

Classification using mutual information approach

The maximum information increase feature selection was used to select a small number of independent features for the target recognition problem. Using this approach the initial class uncertainty may be reduced to zero with a relatively small number of features. These features when used in the combinational classifier will give 100% recognition rate for the training data. In addition, since a small number of features is selected the obtained combinational classifier will have a good recognition rate for the test data. Fig.3 illustrates the total information which can be obtained from a sequence of features selected using MIIFS method. We see that the information content of the selected set of features quickly saturates at 100%. Further increase in the number of features will not add new information about the existing set of training signals. It may, however, increase the robustness of the classifier to recognize test data. This can be compared to solving an over determined set of linear equations. A solution can be obtained with the number of independent equations equal to the number of unknowns (or the rank of the system in case of the rank deficiency), however, considering additional equations gives a more robust solution (considering that the system of equations could have been affected by noise).

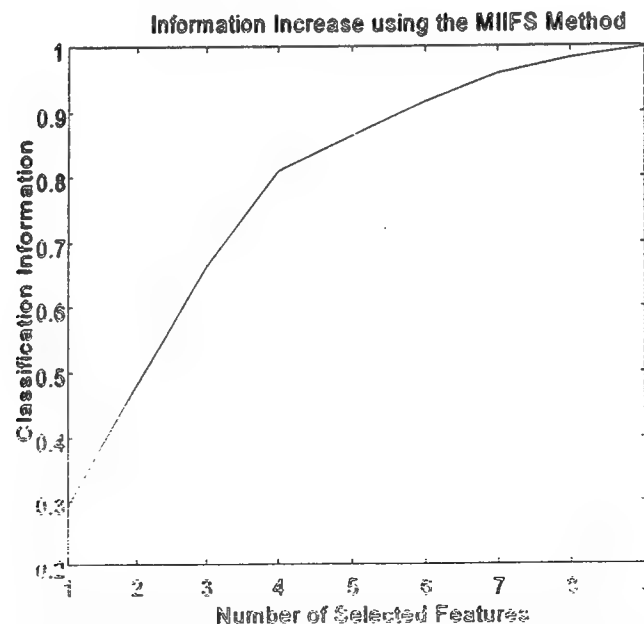


Fig. 3 Information content in a sequence of features.

In order to observe changes in the classification rate using the features which were selected by MIIFS algorithm a number of tests were performed. Fig. 4 shows the changes in the classification rate for targets from all six classes using from 3 to 9 features. We see that the classification rate quickly saturates for each class. This may indicate that classification with a higher recognition rate will be very difficult to achieve (at least in the selected feature space).

Recognition Rate for Information Based Feature Selection Test

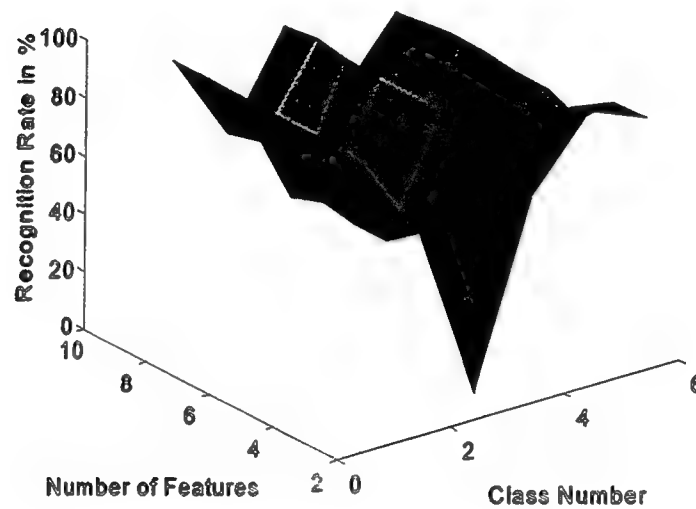


Fig. 4 Change in the classification rate using 3 to 9 features

The result of classification shown in Table 2 corresponds to a combinational classifier which uses 8 features selected by the MIIFS method.

Target number	Training Data		Test Data	
	recognition rate in %	error rate in %	recognition rate in %	error rate in %
1	100	12	93	28
2	81	7	65	23
3	95	12	95	17
4	75	13	65	28
5	95	15	87	17
6	88	7	73	8

Table 2. Classification results using MIIFS method

Conclusion

In this work feature selection in neural networks for high range resolution radar target recognition was investigated. Sequential and combinational classifiers were used as an example of nonorthogonal classifiers to select distinguishing features for synthesis of ontogenic neural networks. An orthogonal classifier based on the dominant distinguishing features was selected to compare the classification performance. Wavelet transforms and other signal transformations were used to preprocess the radar signal data. Simulation results demonstrate a potential benefit of using nonorthogonal classifiers for automatic target recognition.

References

- [1] F. Girosi and T. Poggio, "Networks and the best approximation property," Artificial Intelligence Lab, Memo 1164, MIT (1989).
- [2] T. Chen and H. Chen, "Approximation capability to functions of several variables, nonlinear functionals and operators by radial basis function neural networks," IEEE Trans. Neural Networks, vol. 6, pp. 904-910, 1996.
- [3] Q. Zhao and Z. Bao, "Radar target recognition using a radial basis function neural network," Neural Networks, vol. 9 no. 4, pp. 709-720, 1996.
- [4] D. Ensley and D. Nelson, "Applying cascade correlation to the extrapolation of chaotic time series," Proceedings of the Third Workshop on Neural Networks: Academic, Industrial, NASA, Defense 92; (Auburn AL, Feb. 1992).
- [5] R. Battiti, "Using mutual information for selecting features in supervised neural net learning", IEEE Trans. Neural Networks, vol. 5, pp. 537-550, July 1996.

ON MULTIOBJECTIVE FUNCTION OPTIMIZATION
IN ENGINEERING DESIGN

Alfred G. Striz
Associate Professor
School of Aerospace and Mechanical Engineering

The University of Oklahoma
865 Asp Avenue, Felgar Hall 206
Norman, OK 73019-0601

Final Report for
Summer Faculty Research Program
Wright Laboratory

Sponsored by
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Laboratory

August 1996

ON MULTIOBJECTIVE FUNCTION OPTIMIZATION IN ENGINEERING DESIGN

Alfred G. Striz
Associate Professor
School of Aerospace and Mechanical Engineering
The University of Oklahoma

Abstract

As mathematical optimization methodologies become more and more accepted in the various areas of engineering, the complex problems at hand often require multicriteria or multiobjective function optimizations since most real-life design or decision problems involve multiple and conflicting objectives and constraints. In order to decrease the complexity of such optimizations, it seems of interest to investigate how the various objectives and constraints of a given problem influence and complement each other. Such knowledge could reduce the number of objectives and constraints by eliminating from the optimization loosely coupled parameters. In the present approach, various scenarios were developed, and studies in mathematical, structural, aircraft performance, and aircraft multidisciplinary design optimization were suggested to address these issues. Investigations in structural and aircraft multidisciplinary design optimization were initiated.

Additional MDO issues were addressed during the course of the summer. They are presented in Appendices at the end of the report.

ON MULTIOBJECTIVE FUNCTION OPTIMIZATION IN ENGINEERING DESIGN

Alfred G. Striz

Introduction

Mathematical optimization methodologies have become more and more incorporated into the various areas of engineering design over recent years. Since many complex real-life design and decision problems involve multiple and conflicting objectives and constraints, they often require multicriteria or multiobjective function optimization for their solution. This can result in quite excessive cost (money as well as CPU time). It is, thus, of interest to decrease the complexity of such optimizations. During the course of the summer, the author started to investigate how the various objectives and constraints of a given problem influence and complement each other. Such knowledge could potentially reduce the number of objectives and constraints in a multicriteria optimization by eliminating loosely coupled parameters.

In the present approach, a range of scenarios is developed. Specifically, studies in mathematical, structural, aircraft performance, and aircraft multidisciplinary design optimization are suggested to address these issues. Investigations in structural and aircraft multidisciplinary design optimization have been initiated.

Background

During the course of the summer, the author accumulated background material relating to the issues of interest. The following subject areas were researched:

Multicriteria or Multiobjective Function Optimization - MO

A multicriteria or multiobjective function optimization (MO) problem is defined as an optimization of a system with multiple and often conflicting objectives and constraints. Some of the first applications of MO to problems in structural mechanics, specifically to the explanation of natural structural shapes, were given by Stadler [1]. An extensive treatment of the theory behind MO together with many applications in various fields (approximation theory, welfare economics, resource planning and management, reproductive strategies, aircraft control systems, trusses, focussing systems, etc.) is given by the same author in Reference 2. The subject and its application in the areas of mechanisms and dynamic systems, aircraft and spacecraft structural design, machine tool design, metal forming and cast metal technology, civil and architectural engineering, and structures made from advanced materials were treated in detail by Eschenauer, Koski, and Osyczka (Eds.), in Reference 3.

Edgeworth-Pareto Optimality

A set of design variables is Edgeworth-Pareto [4,5] optimal if the values of the objective functions become worse or do not change due to any small change in any of the design variables. This is an important concept in MO which is sometimes referred to as Pareto optimization.

Game Theory

In the field of business, economics, and politics, game theory has a solid role in strategic planning [6-8]. Specifically, from two-person zero sum games (which are competitive and state that one person can only gain if the other person loses by an equal amount) to n-person non-zero

sum games (which allow for cooperation of the players and are not limited to a fixed resource), many game types are possible. This approach can be translated to engineering design by relating the various objectives and constraints (e.g., minimum stresses and minimum displacements competing for a fixed design weight).

Methodology

The present study focusses on the behavior of the objective function(s) and on the dependencies between the objective function(s) and the constraint(s) of specific optimization problems in order to shed light on the issues of general multiobjective function optimization.

Scenarios

The following optimization scenarios were developed and are being or will be investigated:

1. a mathematical optimization [9]
2. an aircraft performance study (fuel versus payload versus range versus profit)
3. a structural optimization (truss: weight, strength, displacement, natural frequencies) [10]
4. an aircraft MDO study (wing: weight, strength, roll, flutter) [11]

Edgeworth-Pareto Optimal Solutions

For the simple scenarios (mathematical and truss structure optimizations), the results will be compared to Edgeworth-Pareto optimal solutions in order to assure that the optimizers capture global rather than local optima.

Game Theory

The various cases will be likened to n-player zero sum and non-zero sum games. For two player zero sum games, equilibria (saddle points) will be found. The various dependencies in n-player non-zero sum cases will be investigated and compared to the results from optimization studies in the literature.

Methodology

In this investigation, optimizations are performed with interchangeable objective functions and constraint sets for the four scenarios. A general description of the methodology is given in the following for the case of the optimization of a truss structure:

- I. Initially, optimizations are performed with just one objective function and one constraint set (specific example: truss structure where initial objective function = weight, initial constraints = stresses in bars or displacements at nodal points). Possible extrema (saddle points) are established as well as a relationship between the objective function and the constraints.
 1. The structure is optimized for minimum weight with different stress constraints in a given range. No displacement constraints are applied.
 2. The structure is optimized for minimum weight with different displacement constraints in a given range. No stress constraints are applied.
 3. The structure is optimized for minimum stresses with different weight constraints in a given range. The weight constraints are to match the weights obtained in the optimizations in I.1. No displacement constraints are applied.
 4. The structure is optimized for minimum displacements with different weight constraints

in a given range. The weight constraints are to match the weights from the optimizations in

I.2. No stress constraints are applied.

II. Then, optimizations are performed with one objective function and two constraint sets (specific example: truss structure where initial objective function = weight, initial constraints = stresses in bars *and* displacements at nodal points). Possible extrema (saddle points) are established as well as a relation between the objective function and the constraints. One constraint is kept fixed as the other is varied.

1. The structure is optimized for minimum weight with different stress and displacement constraints, each in a given range.
2. The structure is optimized for minimum stresses with different weight and displacement constraints, each in a given range. The weights are to match the weights obtained in the optimizations in II.1. The displacement constraints are to match those used in II.1.
3. The structure is optimized for minimum displacements with different weight and stress constraints, each in a given range. The weights are to match the weights from the optimizations in II.1. The stress constraints are to match those used in II.1.

III. Next, optimizations are performed with two objective functions and one constraint set (specific example: truss structure where initial objective function = weight + stresses, initial constraints = displacements at nodal points). Possible extrema (saddle points) are established as well as a relationship between the objective functions and the constraints.

1. The structure is optimized for minimum weight and minimum stresses with different displacement constraints in a given range.
2. The structure is optimized for minimum weight and minimum displacements with different

stress constraints in a given range.

3. The structure is optimized for minimum stresses and displacements with different weight constraints in a given range. The weights are to match the weights obtained in the optimizations in II.1.

IV. Finally, optimizations are performed with three objective functions (specific example: truss structure where objective function = weight + stresses + displacements)

1. The structure is optimized for minimum weight and minimum stresses and displacements simultaneously.

In all cases, the side constraints (upper and lower bounds) on the design variables stay the same. All objective functions and constraints are scaled as needed. Optimizations I and II provide information to determine the weighting coefficients for the different objective functions used in the multiobjective function optimizations III and IV.

Tools for the Study

The tools for this investigation include: 1) the widely used optimizer NPSOL [12] (based on sequential quadratic programming) for the mathematical optimization scenario and the aircraft performance study; 2) the MDO code ASTROS [13] (which has been made a valuable asset for this research by the addition of a function package, allowing for the user definition of objective functions and constraints) as well as the structural analysis and optimization tool MECHANICA [14] for the structural optimization of the truss; 3) ASTROS for the aircraft MDO study.

Sample Results and ASTROS Function Package Evaluation

Sample results (Cases I.1, I.2, and II.1) obtained for a tenbar truss structure using the function package in ASTROS for the objective function and the constraints showed the strong dependence of the optimal solution (minimum weight) on the displacement constraint and the weak dependence on the stress constraint, indicative of a stiffness driven problem, in Table 1.

Further studies dealt with a displacement objective function (Case II.3) and with a combined displacement/weight objective function (Case III.3). Both of these cases indicated that it is essential to constrain the objective functions (displacement and displacement/weight) in order to reach reasonable optima (see sample results in Table 2).

Also shown in Table 2 are results of an ASTROS function package evaluation, where the built-in objective function weight and the built-in constraints were one by one replaced by a function-defined objective function and by function-defined constraints. All results were quite similar. However, it was found that the function package in ASTROS Version 12 did not allow for an optimization where weight was the only defined function. This case always resulted in an abort at the start of the sensitivity analysis. Only when at least one more function-defined constraint was added did the optimization proceed normally.

A further outcome (and one of the most interesting results of the summer) was that, in minimum weight optimization, various different locally optimal solutions could be obtained

when different weight constraints were put on the (weight) objective function. This seems to imply that it is not always necessary to start with different values for the design variables in order to map the design space and zero in on a global optimum, but that changing constraint(s) imposed on the objective function(s) can fulfil the same purpose. This concept will be tested for other objective functions as well.

References

1. Stadler, W, "Natural Structural Shapes (The Static Case)", *Quarterly Journal of Mechanics and Applied Mathematics*, Vol. 31, 1978, pp. 169-217.
2. Stadler, W., *Multicriteria Optimization in Engineering and in the Sciences*, Plenum Press, New York, New York, 1988.
3. Eschenauer, H., Koski, J., and Osyczka, A., *Multicriteria Design Optimization*, Springer Verlag, New York, New York, 1990.
4. Edgeworth, F.Y., *Mathematical Psychics*, P. Keagan, London, England, 1881.
5. Pareto, V., *Manuale di Economia Politica*, Societa Editrice Libreria, Milano, Italy, 1906.
6. Bennion, E.G., *Elementary Mathematics of Linear Programming and Game Theory*, MSU Business Studies, Michigan State University, Michigan, 1960.
7. Zagare, F.C., *Game Theory: Concepts and Applications*, Sage University Papers Series, No. 07-041, Sage Publications, Newbury Park, 1984.
8. Colman, A.M., *Game Theory and Experimental Games - The Study of Strategic Interaction*, Pergamon Press, Oxford, England, 1982.
9. Venkayya, V.B., Tischler, V.A., Pitrof, S.M., *Benchmarking in Multidisciplinary Optimization*, Report to AIAA MDO TC Benchmarking Subcommittee, April 1995, Wright Laboratory, Wright-Patterson AFB, Ohio, pg. 9 (Problem 18).

10. Haftka, R.T., and Gürdal, Z., *Elements of Structural Optimization*, 3rd Edition, Kluwer Academic Publishers, Boston, Massachusetts, 1993, pg. 237ff.
11. Striz, A.G., Eastep, F.E., and Venkayya, V.B., "Influence of Static and Dynamic Aeroelastic Constraints on Structural Optimization", **AIAA-91-1100-CP, Proceedings, AIAA/ASME/ASCE/AHS/ASC 2nd Structures, Structural Dynamics and Materials Conference**, Baltimore, Maryland, April 1991, pp. 470-476.
12. Gill, P.E., Murray, W., Saunders, M.A., and Wright, M.H., *User's Guide for NPSOL (Version 4.0)*, Technical Report SOL 86-2, January 1986, Stanford University, Stanford, California.
13. Neill, D.J., and Herendeen, D.L., "ASTROS Enhancements, Volume I - ASTROS User's Manual", **WL-TR-93-3025**, Flight Dynamics Directorate, Wright Laboratory, Wright-Patterson Air Force Base, Ohio, March 1993.
14. Enright, C., Glennon, E., and Ring, L., "MECHANICA - Applied Structure Reference Manual", Rasna Corporation, March 1994, San Jose, California.

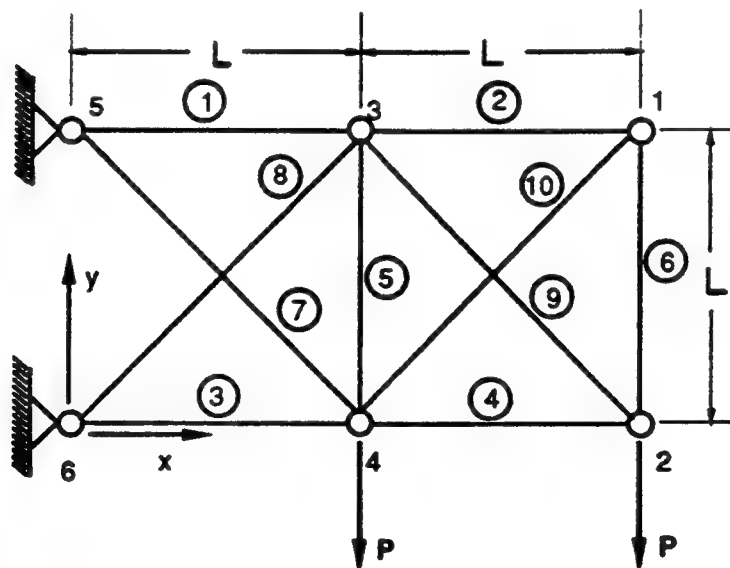


Figure 1. Tenbar Truss with Boundary and Loading Conditions

Table 1. Dependency of Objective Function (Weight in lbs) on Constraints (Material Strength and Allowable Displacement)

Disp.	Stress 10,000 psi	V	C	Stress 25,000 psi	V	C	Stress 35,000 psi	V	C	Stress 50,000 psi	V	C	Stress 75,000 psi	V	C	Stress 1.0E8 psi	V	C
± 1"	10453.2	-	2v	10182.3	1	2v	10195.5	1	2v	10198.2	1	2v	10174.5	1	2v	10185.4	1	2v
± 2"	5393.69	1	1v	5091.39	1	2v	5100.17	1	2v	5102.73	1	2v	5104.57	1	2v	5109.20	1	2v
± 3"	4057.21	-	-	3516.02	1	1v	3411.32	1	2v	3413.15	1	2v	3418.45	1	2v	3401.93	1	2v
± 4"	4050.67	-	-	2674.44	-	1v	2600.29	1	1v	2557.09	1	2v	2557.84	1	2v	2565.22	1	2v
± 5"	4059.18	-	-	2156.14	1	1v	2126.92	1	1v	2097.03	1	2v	2055.32	1	2v	2049.36	2	2v
±∞	4057.21	-	-	1626.31	-	-	1157.71	1	n	813.89	1	n	540.33	1	n	41.97	10	s

V = number of design variables (= bar areas) at the minimum value

C = number of displacement components at the maximum value (v = vertical)

n = linear solution to nonlinear problem (due to large displacements)

s = nonsensical solution due to extremely large displacements

Table 2. Dependency of Final Weight (in lbs) on Choice of Objective Function and Constraint Formulation

Function Obj. Fnct.	Built-in Obj. Fnct.	Function Constraints	Built-in Constraints	Final Weight
Disp. (Case II.3)	-	Disp. [$\pm 1.0''$], Str. [10k], Weight [10455 lbs]	-	10,459.44
Disp. (Case II.3)	-	Disp. [$\pm 2.0''$], Str. [25k], Weight [5092 lbs]	-	5076.57
Disp. (Case II.3)	-	Disp. [$\pm 3.0''$], Str. [35k], Weight [3412 lbs]	-	3391.98
Disp. (Case II.3)	-	Disp. [$\pm 4.0''$], Str. [50k], Weight [2558 lbs]	-	2559.26
Disp. (Case II.3)	-	Disp. [$\pm 5.0''$], Str. [75k], Weight [2057 lbs]	-	2058.03
Disp., Weight (Case III.2)	-	Disp. [$\pm 2.0''$], Str. [25k], Weight [5092 lbs]	-	5098.90
-	Weight	-	Disp. [$\pm 2.0''$], Str. [25k]	5105.33
-	Weight	Disp. [$\pm 2.0''$]	Str. [25k]	5105.79
-	Weight	Str. [25k]	Disp. [$\pm 2.0''$]	5094.96
-	Weight	Disp. [$\pm 2.0''$], Str. [25k]	-	5095.83
Weight	-	-	Disp. [$\pm 2.0''$], Str. [25k]	Error
Weight	-	Weight [5076 lbs]	Disp. [$\pm 2.0''$], Str. [25k]	5077.27
Weight	-	Disp. [$\pm 2.0''$]	Str. [25k]	5104.42
Weight	-	Str. [25k]	Disp. [$\pm 2.0''$]	5105.21
Weight	-	Disp. [$\pm 2.0''$], Str. [25k]	-	5106.51
Weight	-	Disp. [$\pm 2.0''$], Str. [25k] Weight [5076 lbs]	-	5061.32

Str. = Axial Stress Constraints in Truss Bars

Disp. = Displacement Constraints at Nodal Points

APPENDIX A

COMPARISON OF MDO RESULTS FROM ASTROS VERSION 11, VERSION 12, AND VERSION 12/UAI

Problem Statement

The standard F-16 derivative metal wing model with 8,000 lb fuselage half weight (Figure A1) was weight minimized subject to a flutter constraint ($> 25,000$ in/sec), a strength constraint ($< 60,000$ psi tensile and $> -40,000$ psi compressive in the skins during a 9-g pull-up), and a roll effectiveness constraint (> 0.2225). Analyses were performed after the optimizations, both in the discipline order strength - roll - flutter. Only the skins were considered as design variables. With linking, this resulted in a total of 18 design variables, nine on the top and nine on the bottom of the wing.

In the Version 11 runs, two cases were treated:

- Case 1. the PSHELL cards were set to 1.0 inch with the DESVARP initial values reflecting the actual skin thicknesses
- Case 2. the DESVARP initial values were set to 1.0, with the PSHELL cards reflecting the actual skin thicknesses in inches.

In the Version 12 runs, the PSHELL cards were always set to 1.0 inch, with the DESVARP initial values reflecting the actual skin thicknesses (= Case 1).

Comparisons were made between the final weights and the skin thicknesses obtained from ASTROS Version 11, Version 12, and the commercial code, Version 12/UAI.

The total initial weight of the designable part of the wing, i.e., the skins, was 253.49 lbs. The rest of the structural mass and the nonstructural mass added up to a total of 9,844.19 lbs.

Results

The Version 12 and Version 12/UAI results agreed very well. The final weights of the wing skins amounted to 253.81 lbs (Version 12) and 253.11 lbs (Version 12/UAI), an increase of 0.32 lbs and a decrease of 0.38 lbs from the initial weight, respectively.

The Version 11 final designed skin weights for Cases 1 and 2 were 252.52 lbs and 252.64 lbs, respectively. These represented decreases from the original weight of 253.49 lbs by 0.97 lbs and 0.85 lbs, respectively. The differences between the Version 11 and the Version 12 optimized weights were less than 0.51% for all cases.

Additional Results

On various occasions, when the solution sequences for the optimizations or even the analyses in the Version 11 runs were rearranged, DATABASE errors resulted after 7 to 9 iterations in the eigenvalue calculations of the MODES routine preceding the FLUTTER computations, as determined by using the DEBUG MEMORY statement. For these cases, the restart capabilities of ASTROS were employed. This always resulted in converged results which were very close in

magnitude to those from the continuous runs. The maximum V-value on the FLFACT card for all these additional runs was set to 26,416 in/sec as opposed to 30,000 in/sec for all previously mentioned runs.

Specifically, for an optimization sequence of roll - strength - flutter, 252.81 lbs (with upper and lower limits on the design variables left at the values of the initial run) and 252.61 lbs (with upper and lower limits on the design variables adjusted to the restart values of the design variables) were computed for these two different Case 1 runs.

Final designed weights of 252.71 lbs (optimization sequence strength - roll - flutter) and 252.91 lbs (optimization sequence roll - strength - flutter) were obtained for these two different Case 2 runs.

Discussion

The specific set of design variables combined with the specific constraints used here resulted in almost identical initial and optimized weights. However, the distribution of the skin thicknesses changed during the course of the optimization to satisfy the constraints. Specifically, the strength constraint (well satisfied in the skins with maximum/minimum principal stresses of 10,817.8 psi / - 10,879.1 psi, respectively, vs. required maximum tensile and compressive stresses of 60,000 psi and 40,000 psi, respectively) and the flutter constraint (flutter speed 29,753 in/sec vs. a required flutter speed $> 25,000$ in/sec) were already satisfied by the initial design, whereas the roll constraint was not (initial roll effectiveness = 0.1275 vs. required roll effectiveness = 0.225). After

the optimization, the analyses of the skins showed maximum/minimum principal stresses of 20,138.6 psi / -20,032.2 psi, respectively, a flutter speed much larger than 30,000 in/sec, and a roll effectiveness of 0.2683. No minimum or maximum gage sizes were reached.

During these studies, the above mentioned MEMORY DEBUG option as well as the MAPOL language option were exercised, the latter to print a MAPOL solution sequence for ASTROS.

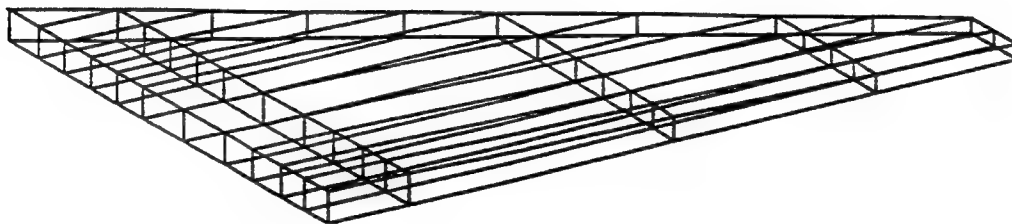


Figure A1. Fighter Wing Finite Element Model

APPENDIX B

PAPER PREPARATION FOR 6TH AIAA/NASA/USAF/ISSMO MULTIDISCIPLINARY ANALYSIS AND OPTIMIZATION SYMPOSIUM

A paper was prepared by the author during the course of the summer with the title

"Comparative Evaluation of Two MDO Codes in Aircraft Wing Analysis and Optimization"

for presentation at the 6th AIAA/NASA/USAF/ISSMO Multidisciplinary Analysis and Optimization Symposium in Bellevue, Washington, September 4-6, 1996. In this paper, results are presented from an aircraft wing MDO study by a M.S. student at the University of Oklahoma, Mr. Shuming Yan.

Abstract

The overall objective of the project was to compare the capabilities and performance of two large scale finite element codes, ASTROS and MSC/NASTRAN, in the areas of finite element and aeroelastic analysis, single constraint optimization, and multidisciplinary optimization with aeroelastic constraints. The low aspect ratio fighter type wing from Reference 11 was selected and modeled by two fully built-up finite element models, one a rather coarse conceptual design model, the other a more complex preliminary design model. First, static, dynamic, and aeroelastic analyses were performed to test the capabilities of the two codes in providing relevant information for the optimization procedure. Then, single constraint optimizations and multidisciplinary optimizations were conducted, using such constraint disciplines as static strength, natural frequency, flutter, and aileron effectiveness, to investigate the optimization performance of the codes.

The results showed that ASTROS and MSC/NASTRAN were, in general, in good to excellent agreement for the disciplines of strength, modal analysis, and flutter in both, analysis and optimization. However, they disagreed to a considerably larger extent in the discipline of aileron effectiveness. For these cases, ASTROS seemed to have some problems with the steady aerodynamic analysis routine (a derivative of USSAERO) which affected the convergence to an optimum. On the other hand, MSC/NASTRAN does not, at present, allow for a simultaneous symmetric flutter and anti-symmetric roll optimization using a half model. This resulted in longer CPU times for the code when a more complex semi-MDO optimization procedure was used in MSC/NASTRAN. When comparing results between the coarse and the complex wing models, the complex model, although it had initially almost the same weight and similar free vibration characteristics as the coarse model, exhibited a lower distributed stiffness in torsion, resulting in considerable differences in the roll effectiveness between the two models. These differences were magnified in the optimizations. For all cases, CPU time and number of iterations were determined and compared. As expected from a highly performance optimized analysis code, MSC/NASTRAN was faster in the analysis applications. In the optimizations, the performance advantage depended on the specific case. It was suggested that additional subsonic steady aerodynamic capabilities be incorporated into ASTROS (presently underway) and/or that the present USSAERO implementation be further examined through the comparison of pressure data with those from other codes. Also, additional comparative investigations were recommended, including the multidisciplinary optimization of a full wing model by MSC/NASTRAN, and the optimization of wing models with lighter or no design variable linking by both codes. Finally, full airplane configurations including tail and fuselage should be investigated for both wing models and both codes.

APPENDIX C

ABSTRACT PREPARATION FOR 38TH AIAA/ASME/ASCE/AHS/ASC STRUCTURES, STRUCTURAL DYNAMICS, AND MATERIALS CONFERENCE

An extended abstract was prepared at the end of the summer program with the title

"Issues of Multiobjective Function Optimization in Engineering Design"

by the author and his lab focal point, Dr. V.B. Venkayya of WL/FIBAD, for the 38th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Kissimmee, Florida, April 1997. In this paper, results are to be presented from the study initiated this summer.

Abstract

The design of complex systems in various areas of engineering often requires multicriteria or multiobjective function optimizations since most real-life design or decision problems involve multiple and conflicting objectives and constraints. In recent years, this process can be aided more and more by mathematical optimization methodologies, limited only by the complexity of the given problem. In order to decrease this complexity, an investigation is suggested to determine how the various objectives and constraints of a given problem influence and complement each other. This can result in simplifications to the optimization procedure by eliminating loosely coupled parameters. Thus, the present study focusses on the behavior of the objective function(s) and on the dependencies between the objective function(s) and the constraint(s) of specific optimization problems. Various scenarios in mathematical, structural, aircraft performance, and aircraft multidisciplinary design optimization are developed and investigated to address these issues.

OPTICAL AND ELECTRO-OPTICAL STUDIES OF POLYMERS

Barney E. Taylor
Visiting Assistant Professor
Department of Physics

Miami University - Hamilton
1601 Peck Blvd.
Hamilton, OH 45011

Final Report for:
Summer Faculty Research Program
Wright Laboratory
WL/MLBP

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

and

Wright Laboratory

September 1996

OPTICAL AND ELECTRO-OPTICAL STUDIES OF POLYMERS

Abstract

Barney E. Taylor
Visiting Assistant Professor
Department of Physics
Miami University - Hamilton

The Physics Group routinely performs room temperature absorption and photoluminescence (PL) measurements on polymers synthesized within the Polymer Branch of Wright Laboratories. These routine measurements were supplemented by a temperature dependence study of the PL of the polymer 6F-PBO over the temperature range of 110 to 365 K. 6F-PBO is a member of the polybenzoxazoles with 6F moieties in the form of alkoxy pendants ($\text{OC}_{10}\text{H}_{21}$). The results indicate a thermally activated quenching of the PL with an activation energy of 109 meV. The individual PL spectra were deconvoluted by means of nonlinear least square analysis. Three gaussian peaks provided an excellent fit to the spectra at all temperatures. Attempts to perform complementary optical absorption measurements in the band-tail region over a similar temperature range were inconclusive due to experimental difficulties that have been identified, but have not yet been overcome.

OPTICAL AND ELECTRO-OPTICAL STUDIES OF POLYMERS

Barney E. Taylor

Introduction

The physics group of the Polymer Branch of WL/MLBP supports the synthetic chemists by electrically and optically evaluating the prototype polymers in order to meet the needs of the Air Force. Ongoing efforts in the physics group is the fabrication and characterization of electroluminescent (EL) devices. In an attempt to understand the physics involved, complimentary studies have also been a part of the continuing studies. Optical absorption and photoluminescence (PL) data are routinely taken on polymers of interest as potential EL materials. Previous efforts of this investigator have been directed toward characterization of EL devices with Poly (p-phenylenebenzo-bisthiazole) (PBZT) [1, 2], and the subsequent PL characterization of PBZT [3].

This work reports on the efforts to further refine the analysis of the PBZT data and new studies on the optical absorption, PL, and contributions to the EL study of a new candidate for viable EL devices - the polymer 6F-PBO. 6F-PBO is one member of the polybenzoxazoles with 6F moieties in the form of alkoxy pendants ($\text{OC}_{10}\text{H}_{21}$). A previous photoluminescence survey of a copolymer containing 6F-PBO yielded a strong PL emission with relatively blue spectral peak[3], hence, it was decided to further investigate the optical properties of 6F-PBO. During the course of the SFRP program, a NRC Fellow joined the physics group, and began fabricating and characterizing EL devices with 6F-PBO as the active layer.

Contributions were made to many different activities within the Physics group during the tenure of the SFRP program: re-visiting the analysis of the temperature study of the PL of PBZT in light of improved calibration techniques; measurement of the UV-VIS-IR absorption spectrum of 6F-PBO samples as a function of sample temperature and analysis of the absorption data in the visible region; PL studies of 6F-PBO as a function of temperature and analysis of that data; design of a mounting assembly to allow routine measurement of absorption spectra as a function of temperature; and development of programs to assist in the acquisition of current versus voltage as well as current and emitted intensity versus voltage data for prototype EL devices. Minor

contributions were made in the areas of maintaining and updating the literature database on EL devices and materials and in developing a database of previously acquired UV-VIS-IR spectra. Each of the topics will be discussed later.

EXPERIMENTAL

Success in the temperature dependent PL measurements during the 1995 SFRP lead to the desire to perform temperature dependent absorption studies. One area of interest is the so called band-tail (the region of energies slightly below the absorption edge). A number of models have been developed for the temperature behavior of this region for other physical systems [4], so it was decided to see if polymeric samples behaved analogously. The Hitachi UV-VIS-IR spectrometer has a closed sample chamber, with a light tight lid. Hence, a means of mounting the cryostat in the optical path of the UV-VIS-IR spectrometer was needed. The lid was removed, a makeshift lid fashioned out of heavy cardboard, and black polyethylene sheeting. The cryostat was placed in the optical beam of the spectrometer by means of a small lab jack. An aperture was placed in the reference beam with the same diameter as the unmasked region of the sample. A series of absorption measurements were made at temperature ranging from 110 K to 475 K. Numerous experimental problems were encountered - most dealt with maintaining light tightness and the position of the cryostat in the beam of the spectrometer while refilling the cold finger of the cryostat. Frosting of the windows of the cryostat was another problem, although the window heaters were active, some frosting still occurred. However, data was taken on a 6F-PBO sample that had been previously characterized by room temperature UV-VIS-IR and PL methods. The data exhibited a temperature dependence in the transmittance. The data will be discussed in a later section.

The experimental difficulties in performing temperature dependent absorption studies lead to the design of a new mounting plate to rigidly hold the cryostat in position. In addition, a second lid for the sample chamber of the spectrometer was ordered, so that the lid could be drilled to allow the cryostat to be inserted. Once the cryostat is positioned, a strong purge of dry nitrogen gas will be used to prevent the windows from frosting. The

increased rigidity will allow the liquid nitrogen reservoir of the cryostat to be refilled without altering the optical alignment. Hopefully, the new mounting system will allow data to be obtained with far less experimental variation, resulting in the ability to perform more conclusive studies of the band-tailregion of polymers.

In light of the experimental difficulties in performing temperature dependent studies in the UV-VIS-IR spectrometer, it was decided to attempt a set of single beam temperature dependent studies using the components of the PL apparatus. The OMA (optical multichannel analyzer) was set to respond to the visible range near the absorption edge of 6F-PBO. The monochromator on the tungsten lamp (normally used to excite photoluminescence) was set to zeroth order so that light of all wavelengths would be emitted. A means of rigidly mounting the cryostat to the optical table was devised. The measurement process was as follows: 1) An aperture the same size as the open sample area of the cryostat was placed at the sample location, and a baseline transmittance spectrum was measured using the OMA. 2) The cryostat was evacuated and placed into its mount. A spectrum altered by the presence of the windows of the cryostat was then obtained. 3) The 6F-PBO sample was installed, and the cryostat was re-evacuated and subsequently mounted. A room temperature spectrum was measured. Liquid nitrogen was added, and the sample was allowed to cool to 110 K. Upon reaching thermal equilibrium as indicated by the temperature not changing for at least 5 minutes, a spectrum was taken. The set point of the temperature controller was adjusted up a few degrees, and the process repeated. Spectra were obtained at temperatures between 110 K and 365 K. Background spectra were taken before and after the measurements for use in analyzing the data. As was expected, no change in the background was observed during the course of the experiments. A normalization spectra of the calibrated lamp were also obtained without disturbing the OMA's position or settings. The analysis of the data will be discussed in the next section.

Dr. Gang Du, a NRC Fellow, joined the physics group in June, and began a program of fabricating and characterizing polymeric EL devices. Based upon the relatively high PL brightness and the blue nature of the PL spectrum of the 6F-PBO copolymer measured during the 1995 SFRP[3], he began a study of 6F-PBO as the active layer of EL devices. He has been quite successful in achieving luminescence, but the devices have been short lived. In support of

Dr. Du's effort, a temperature dependent study of the PL of 6F-PBO was performed over the temperature range of 110 K to 365 K. A methodology similar to that developed for the 1995 SFRP was used to obtain the data. New filters had been acquired, enhancing the ability to take PL data without artifacts due to the experimental process. A new sample of 6F-PBO was cut from the as-cast film, and mounted into the cryostat. The experimental process was not much different than that just described for absorbance. After alignment using the zeroth order beam of the monochromator, the excitation wavelength was chosen, a room temperature reference PL spectrum was taken. Next, the liquid nitrogen reservoir was filled, and upon reaching equilibrium, the lowest temperature spectrum was taken. The temperature controller's set point was altered, and the process repeated for temperatures up to about 365 K. Frosting of the window occurred at about 200 K when the nitrogen reservoir was unintentionally overfilled. A change in the PL spectrum was noticed - an extra peak, that eventually subsided. On the following day, a strong stream of dry nitrogen was directed across the window, and the measurement repeated. The room temperature and lowest temperature measurements were nearly identical. The anomalous peak was never observed. The second set of data was later normalized and analyzed, and the results will be discussed later.

As an aid to Dr. Du's effort on fabricating and characterizing EL diodes, programs were developed to acquire voltage versus current data on the EL diodes using a Keithley 237 Source Measurement Unit connected to a computer by the GPIB bus. Two programs were developed. One, a DOS program, was very primitive and not easy to use in the subdued light associated the pre-EL measurements. The program was rewritten as a windows program so that many functions could be performed by clicking on the desired choice. These programs communicated with the Keithley 237, instructing it to set the current (or voltage) and to measure the complementary variable. The resulting data was saved into a disk file. During the earlier EL studies of PBZT [1], it was found that a simultaneous measurement of optical 'turn on' while acquiring the I-V data was very useful. The windows program was modified to perform such measurements by including some of the instrumentation used for EL studies before the arrival of the OMA. The new program was developed to allow simultaneous measurement of the Voltage versus current and Intensity versus current for the EL diodes. The Keithley 237 was augmented by a PAR 5210 Lock

In Amplifier. The PAR 5210 was used as a GPIB A/D converter to record the signal level from the optical detector. The optical detector can be either a large area Hamamatsu silicon diode, or a photomultiplier. The current from either is fed into a PAR 181 Current to Voltage Preamplifier to obtain a measurable dc voltage level. The history of polymeric EL devices in this lab has been a strong variation in intensity over short periods of time. Hence, the program included the ability to repetitively measure the voltage and intensity at a given current, and to calculate and report the mean and standard deviation. Stability of the operating device can be judged by observing the size of the standard deviation at the various measurement points.

A RDL summer student had, as part of his project, the task of setting up a database of literature articles related to luminescence. Some assistance was provided to the student in setting up and using a program designed for such databases by a previous RDL summer student.

Data and Discussion:

Optical Absorption of 6F-PBO: Although much time and energy were invested in the attempts to measure meaningful data from the band-tail region of the visible absorption spectrum, no definite results were found. The present situation with the dual beam Hitachi UV-VIS-IR spectrometer precluded having a rigid, light tight mount for the cryostat. With the single beam OMA measurements the measured dynamic range was not sufficient to duplicate the room temperature spectrum of the Hitachi, except for the region near the band edge. It is likely that the limit to the dynamic range of the OMA is the stray light within the laboratory. The seals to the door were not completely effective. Also, numerous devices have indicator lights, LCD displays and so forth. Once the eye is dark adapted, the shape of a person wearing light colored clothing can be discerned as they move. This level of light pollution has not been a problem with the PL measurements, but it likely reduced the available dynamic range for making useful single beam absorption measurements on 6F-PBO using the OMA.

The fact that data was taken over a range of temperatures from 110 K to 475 K in the Hitachi indicates that with a suitably rigid and repeatable

mounting system and good light baffling, meaningful studies of the absorption of polymeric samples can be performed over a temperature range of sufficient span to allow investigation of the band-tail as a function of temperature. Upon completion of the fabrication of the mounting plate, and the modification of the new lid, the experiment can be completed.

One interesting phenomena was observed during the investigation of the absorption as a function of temperature. During the measurements the 6F-PBO film lost material. This was indicated by a film formed on the cryostat window that was approximately 7 microns thick. The thickness of the film was deduced from analysis of the interference fringes in the spectrum of the empty cryostat (after removing the sample). The loss in thickness was about 40%. The initial film thickness was about 110 microns, and the post heating thickness was about 70 microns. Thermal Gravimetric Analysis (TGA) was performed on a previously unused sample of the material in a helium atmosphere. The study indicated no loss of mass before temperatures of about 400 °C. An Isothermal anneal at 190 °C (comparable to the highest temperatures measured in the cryostat) revealed a loss of mass of only 0.68% over a 5 hour interval [5]. One could possibly think of the relaxation of the polymeric chains leading to the reduction in the thickness of the sample, if it were not for the film deposited on the cryostat window. The Transmittance spectrum of the sample in air, the sample in the cryostat before heating, and the sample in air after heating were parallel. The change in thickness of the sample caused a greater transmittance after heating. The transmittance of the windows plus deposited film exhibited the same shape as the bulk film with superposed interference fringes due to its small thickness. Due to a shortage of the 6F-PBO material in film form, the experiment was not repeated. Subsequent studies as a function of temperature were done at temperatures less than 100 °C, and no mass loss was noticed. This is consistent with an earlier report of T_g for this material [6].

PBZT Photoluminescence: During the 1995 SFRP program, this investigator performed a series of temperature dependent PL measurements on the rigid rod polymer PBZT. The standard calibration lamp needed to properly interpret the spectra arrived very close to the end of the tenure. Hence, the accuracy of the spectral intensity normalization of the previously reported PL data was

somewhat in doubt. During the course of the year, improvements had been made in the techniques used to normalize PL data. These new methods were used to reanalyze the earlier PBZT PL data. A computer program substantially written by Max Alexander (SOCHE Program) was enhanced to apply the refined calibration technique to the existing PBZT PL data. Very slight changes were noticed in the normalized spectra - primarily in the extremities of the spectrum where the measured intensity of the signal is small. The normalized area was studied as a function of temperature, and each spectrum was analyzed separately by means of non-linear least squares fitting of three or more gaussian functions to the spectra. The results of the fits were quite similar to those obtained during the 1995-SFRP on the less well normalized data. Only slight differences were encountered. The normalized area versus temperature graph is shown in Figure 1. The data points are indicated by the circles and the best fit line is the solid curve. Pankove [7] describes the process as a reduction in the quantum efficiency as the temperature increases, due to thermally activated nonradiative recombination sites. The integrated area and the temperature should obey the following relationship:

$$\eta(T) = \frac{1}{1 + C e^{-E^*/kT}} = \frac{PL(T)}{PL(0)} ,$$

where η is the quantum efficiency and E^* is the thermal activation energy. Since the $T = 0$ K PL must be known to calculate η , a slightly modified form was adopted for nonlinear least squares fitting:

$$PL_T = \frac{A}{1 + C e^{-\frac{T_0}{T}}} ,$$

where A is effectively the integrated area at $T = 0$, and T_0 is the equivalent temperature of the thermal energy. Hence, the activation energy is given by $E^* = k_b T_0$. The activation energy obtained from the nonlinear least squares fit is 84.5 meV which is exceedingly close to the 83 meV obtained in the 1995 analysis of this data. The results of fitting multiple gaussian or lorentzian lineshapes to the data is substantially identical to those obtained in the 1995 SFRP study [3]. Hence they will not be repeated here.

6F-PBO Luminescence: The PL data for the 6F-PBO temperature dependent study were very nicely behaved. Figure 2 is a composite of the normalized

spectra taken at the various temperatures. The largest PL intensity was observed at the lowest temperature, and the PL intensity decreased monotonically as the temperature was increased. Non-linear least squares analysis was performed on each spectrum. Only three gaussian peak were necessary to duplicate the experimental data for any of the spectra. Figure 3 is a typical result of the fitting process. The data is marked by circles, the overall best fit line by the solid line, and the individual peaks that comprise the solid line are shown by the dashed lines. Figure 4 is a plot of the normalized area versus the inverse temperature for the 6F-PBO data. The data was fit to the Pankove function mentioned earlier. The activation energy of the nonradiative processes was found to be 109 meV, about 20% higher than that of the PBZT previously discussed. The coefficients were plotted individually as a functions of temperature and inverse temperature, and the graphs inspected for Arrhenius or other behavior. Several interesting trends were observed, although it is only possible to speculate as to the causes. Figure 5 shows that the energy of each of the three constituent peaks of Figure 3 decreased in energy with increasing temperature from 110 K to room temperature. Above room temperature, the two major peaks increased in energy with increasing temperature while the smaller central peak decreased at a faster rate than at lower temperatures. The behavior below room temperature is observed in semiconductors such as GaAs. The decrease in the energy is tied to the expansion of the lattice as the temperature increases. It does not seem unreasonable to attribute the low temperature trends in the energy to a thermal expansion of the 6F-PBO. The behavior above room temperature is more puzzling. A simple thermal expansion would keep the energy decreasing. At this time no mechanism is known for this unusual behavior. The widths of the gaussians show a linear increase with temperature for all three constituent peaks. The higher energy peak has the greatest broadening with a slope of $7.5 \cdot 10^{-5}$ eV/K. The smallest, central peak has the least increase as the temperature is increased at $1.4 \cdot 10^{-5}$ eV/K. The amplitude of the smaller, central constituent peak was described by either a linear decrease with temperature or by the Pankove expression. The amplitude of the two larger components exhibited significant curvature when plotted as a function of temperature, and were well fit only by the Pankove equation for nonradiative processes. This is not surprising, since the width of the peaks varies very

slightly, hence, the amplitude is almost proportional to the area of the peak. Figure 6 shows the behavior of the constituent amplitudes. The activation energy for each of the constituent peaks is different. Instead of 109 meV, the individual energies range from 71.4 to 123.2 meV. An average using terms weighted by the amplitude would likely result in observed activation energy.

SUMMARY AND SUGGESTIONS FOR FURTHER STUDY

The temperature dependence of the band-tail is potentially a very rewarding study on these systems. Tauc [4] has a very nice explanation of the regions found in typical absorption band-tails in amorphous materials. There should be three distinct regions as one moves to lower energy from the band edge. These regions are the strong absorption region, an exponential region and a weakly absorbing tail. The strong absorption region is the type of absorption discussed in introductory solid state texts, where the absorption coefficient is given by $abs \propto \sqrt{h\nu - E_g}$. The exponential region is described by $abs \propto \exp(-\frac{h\nu}{E_t})$. The energy E_t has been found to be nearly temperature independent in glasses below room temperature, and to decrease slightly with temperature above room temperature. The third region is the Urbach Edge [8], where the sample is found to be weakly absorbing relatively far below the band edge of the material. The nature of the tail can be highly temperature dependent. Preliminary analysis of the Hitachi room temperature data on 6F-PBO indicate that three discernible regions are present. The questionable temperature dependent data did indicate a change in the transmittance with temperature, but the results were inconclusive. Hence, it is expected that the temperature dependent absorption studies will prove useful to the Physics group at a later date.

The results of reanalyzing the existing PL data on PBZT were very satisfying. The similarity in the values of the activation energy clearly indicates that the original calibration, which was thought to be questionable, was, actually, pretty good. Since PBZT does not appear to be a viable candidate for EL devices, it is unlikely that future work will be performed on the PL of PBZT by the Physics Group.

6F-PBO is a promising material at the present time. Dr. Du is fabricating devices that exhibit electroluminescence much more frequently than the PBZT based devices previously investigated, although the device lifetime is very short [9]. Hopefully, the study of the PL of 6F-PBO as a function of temperature will assist in the understanding of the physics of 6F-PBO. The family of PL curves at different temperatures is nicely monotonic. The integrated area versus inverse temperature plot of Figure 4 agrees very well with Pankove's account of nonradiative processes being thermally activated to quench the PL. The activation energy is roughly comparable to that obtained for PBZT and seems reasonable. The ability to fit the spectrum with only three gaussians was a pleasant surprise. In the case of PBZT, the number of peaks needed to adequately describe the spectrum varied at the temperature extremes. For 6F-PBO, three peaks were quite adequate for all temperatures measured. The trend in the width of the individual peaks, to broaden as the temperature increases, seems reasonable in light of the greater thermal energy available. The behavior of energy as a function of temperature is quite surprising above room temperature. From an elementary viewpoint, one would expect the energy to continue to decrease until reaching some type of phase transition. The tendency of the energy to increase with increasing temperature above room temperature is very interesting and should be studied more.

Several questions have been answered, but many more exist for the Physics Group to answer in understanding the optical and electro-optical behavior of their prototype polymers.

REFERENCES

- 1 B. E. Taylor, Final Report, AFOSR SFRP Program, Sept., 1994.
- 2 J. B. Ferguson, R. J. Spry, B. E. Taylor, S. J. Bai, Bull. Am. Phys.
Soc. 22, ??? (1995).
- 3 B. E. Taylor, Final Report, AFOSR SFRP Program, Sept., 1995.
- 4 Jan Tauc, *Amorphous and Liquid Semiconductors*, Chapter 4, Plenum Press,
London, 1974.
- 5 Marlene Houtz, Private Communication.
- 6 Lisa R. Denny, Robert C. Evers, Bruce A. Reinhardt, Marilyn R. Unroe, My
Dotrong, and Marlene Houtz; 22nd International SAMPE Technical
Conference, 186, (1990).
- 7 J. Pankove, Optical Processes in Semiconductors, p. 166, Dover, 1971.
- 8 Franz Urbach, Phys. Rev. 92, 1324 (1953).
- 9 Gang Du, Private Communication.

Integrated Area vs 1/T for PBZT

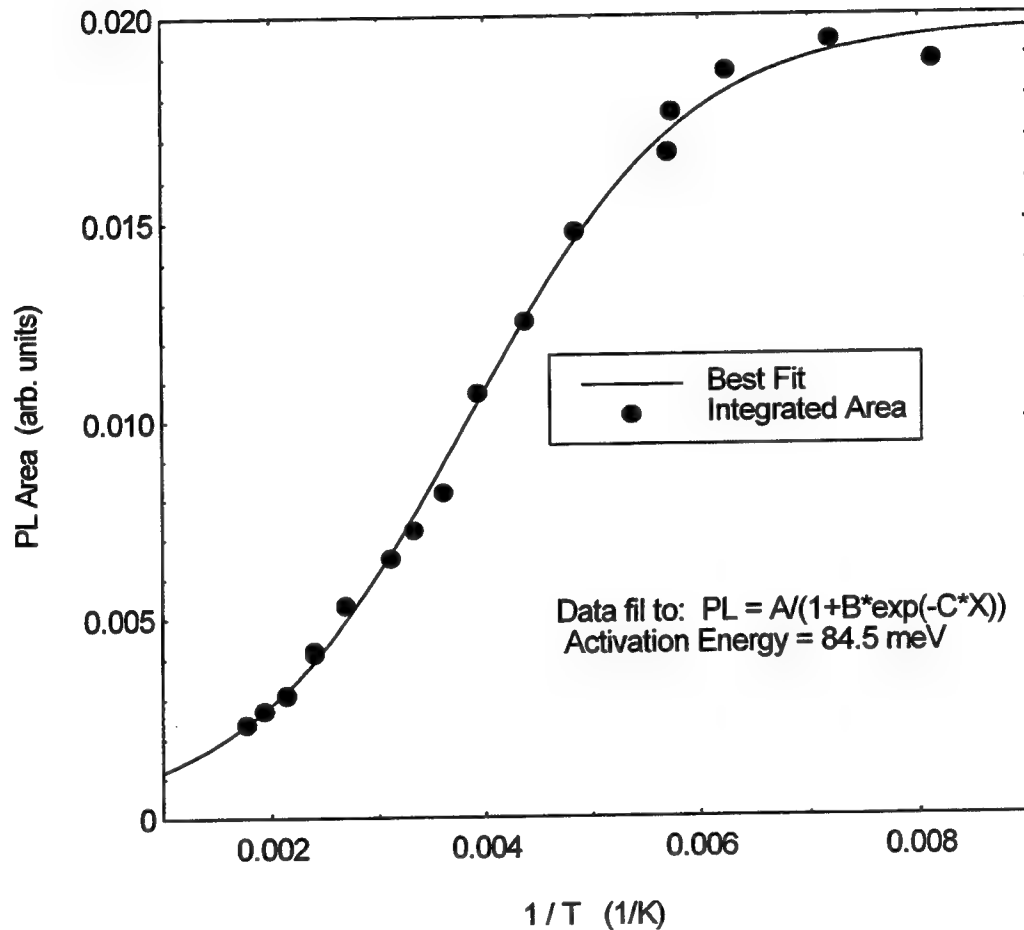


Figure 1. Integrated area of PBZT PL versus inverse temperature. The data was obtained during the 1995 SFRP and was reanalyzed in light of a refined technique for spectral normalization of the raw data. The graph indicates a thermally activated mechanism for the quenching of PL with an activation energy of 109 meV.

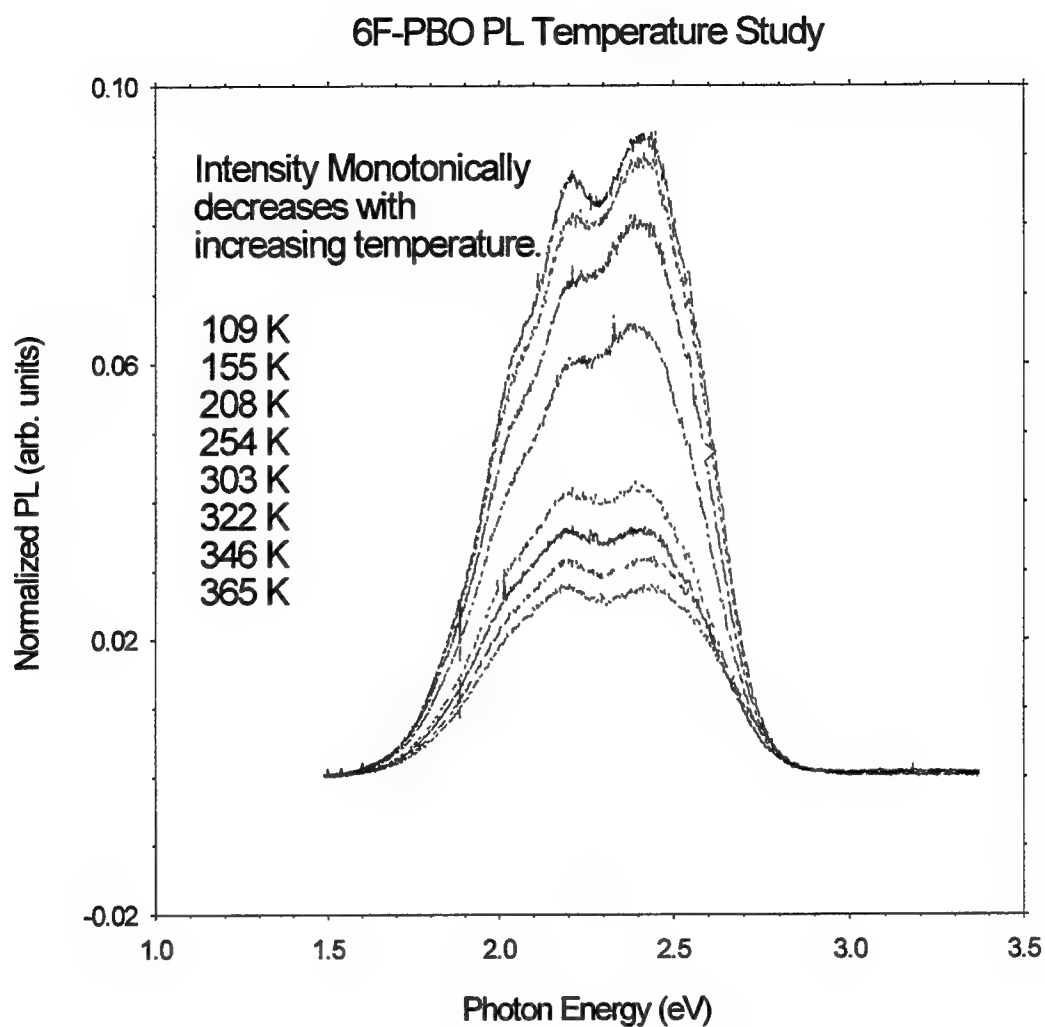


Figure 2. The normalized PL spectra for 6F-PBO is shown as a family for temperatures between 109 and 365 K. The intensity decreases monotonically with temperature.

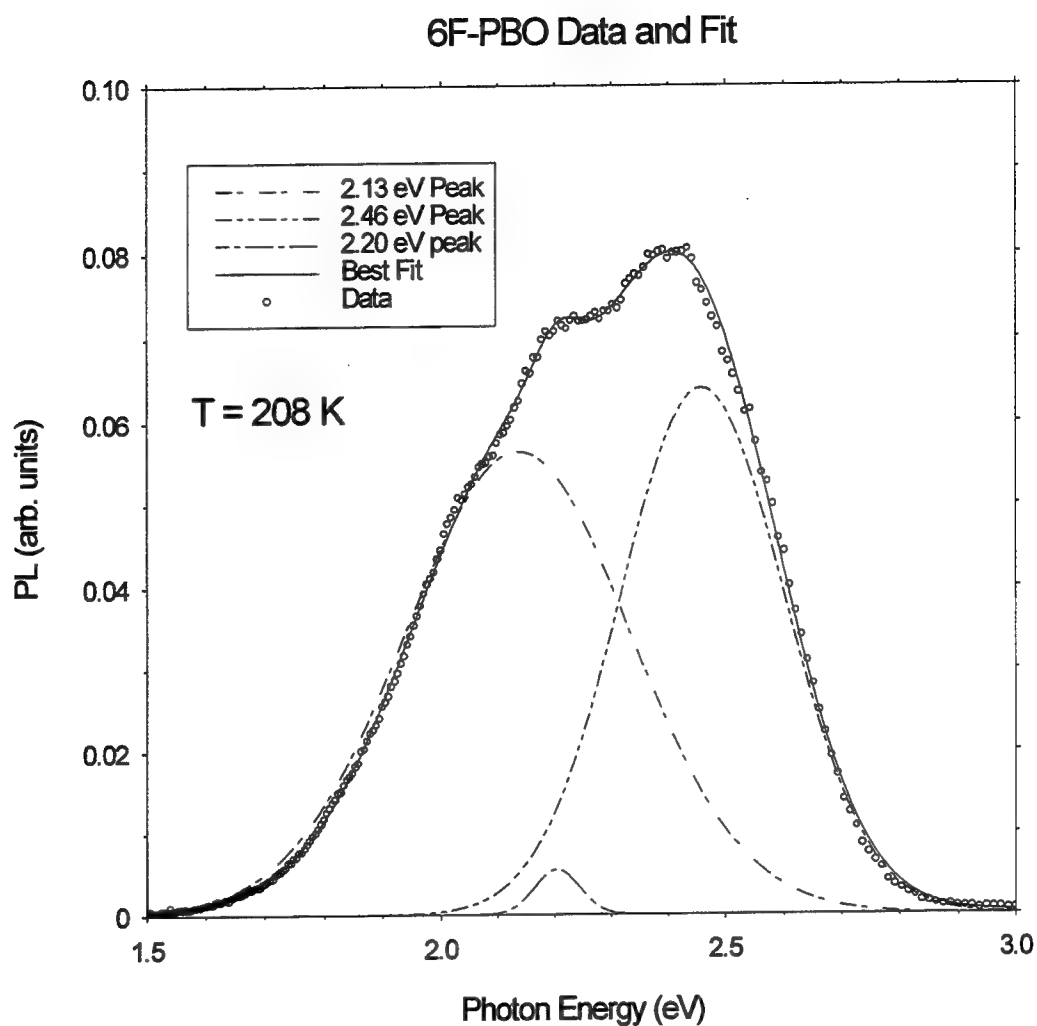


Figure 3. This graph is typical of those obtained by fitting 3 gaussian peaks to the measured photoluminescence spectra. The symbols represent the data, the solid line the overall fit, and the dashed lines the individual peaks. Each peak is characterized by a center energy, a width, and an amplitude.

Integrated PL Area vs 1 / T for 6F-PBO

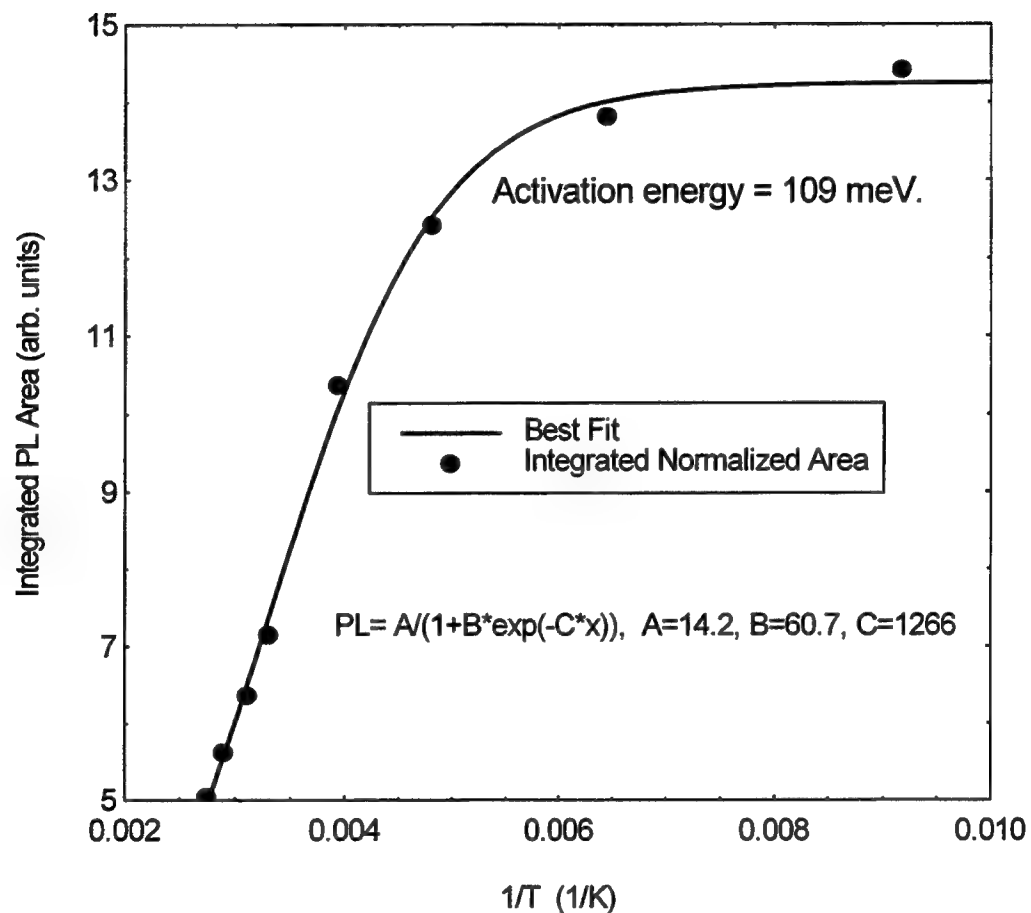


Figure 4. The integrated area versus inverse temperature for 6F_PBO PL. The data has been fit to the Pankove function, resulting in an activation energy of 109 meV.

Peak Energy vs Temperature for 6f-PBO

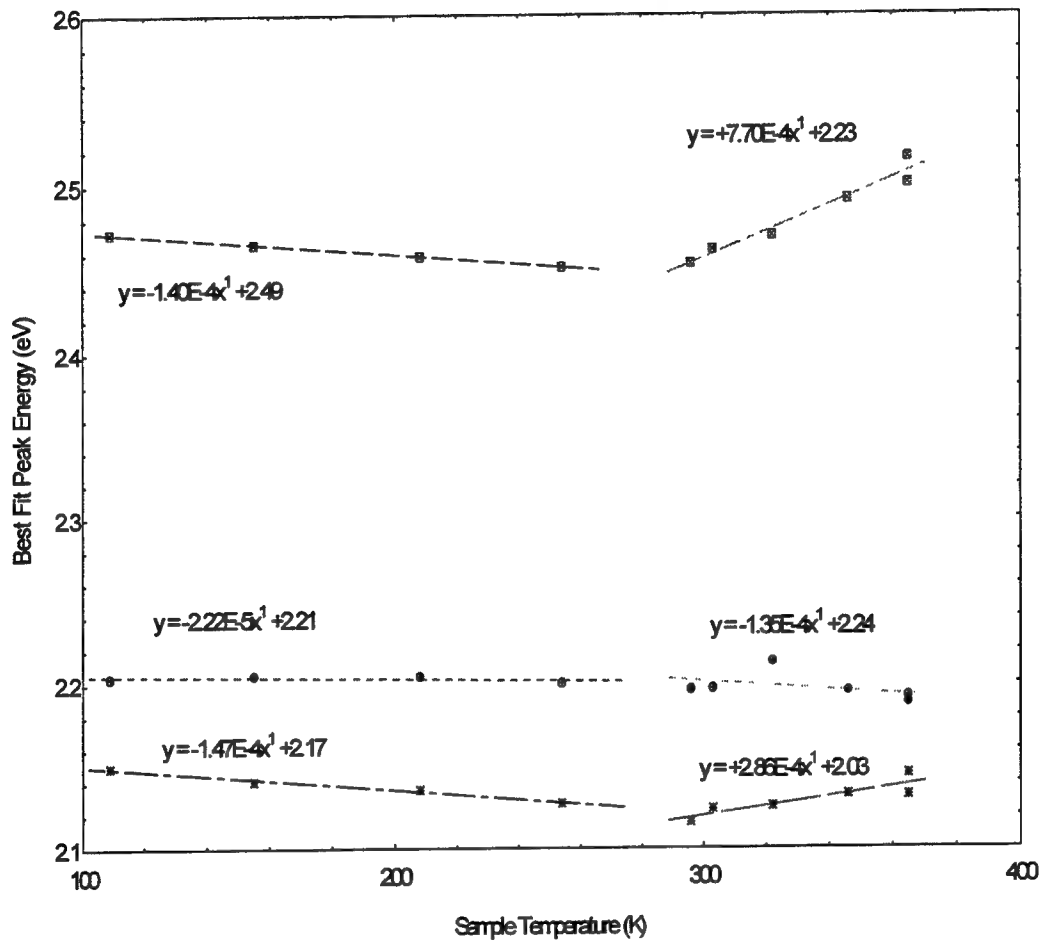


Figure 5. A plot of the best fit center energies versus temperature shows two trends. All of the energies decrease as the temperature rises toward room temperature - as might be expected due to thermal expansion. Two of the peaks increase in activation above room temperature.

Amplitude vs 1/T for 6F-PBO Fits

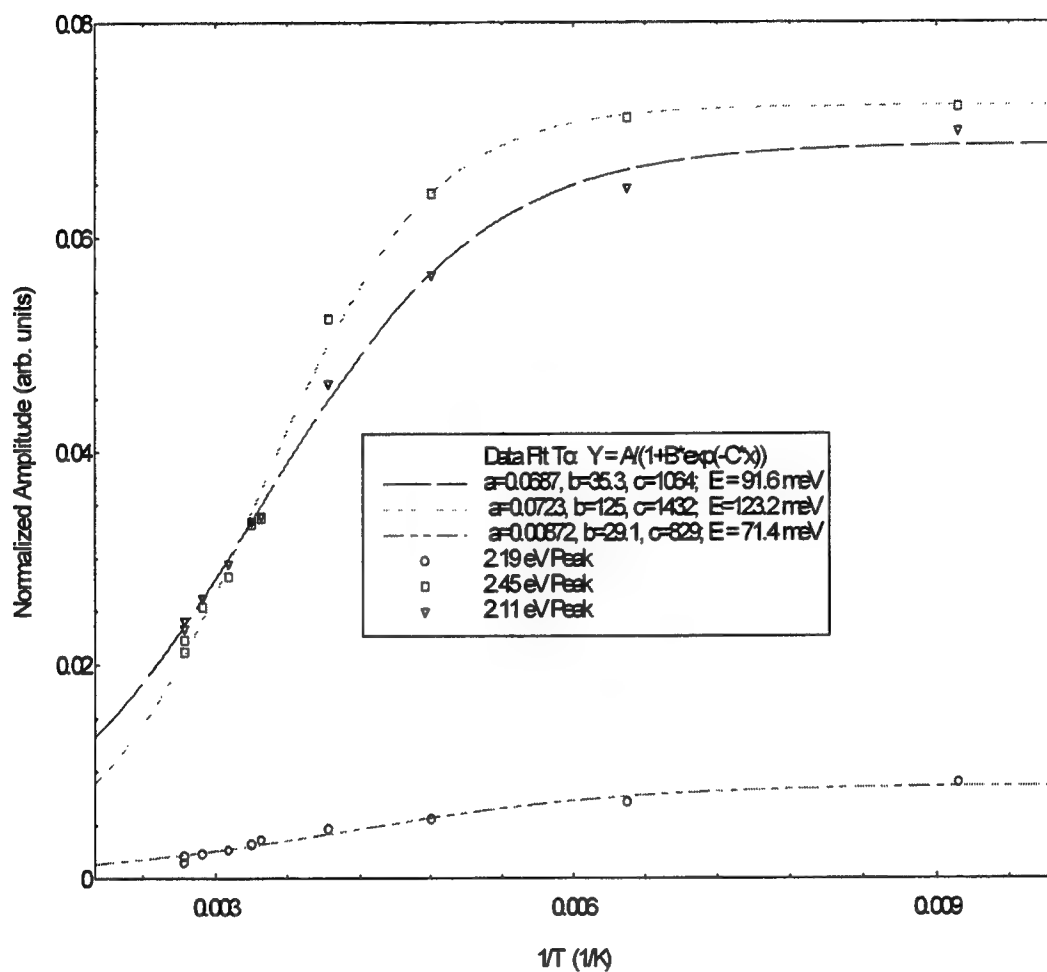


Figure 6. Is a plot of the amplitudes of the individual peaks versus 1/T for the 6F-PBO spectra. Each peak shows a thermal deactivation, with slightly different activation energies.

EFFECTS OF AIRBLAST CHARACTERISTICS ON
STRUCTURAL RESPONSE

Joseph W. Tedesco
Gottlieb Professor
Department of Civil Engineering

Auburn University
Auburn AL 36849

Final Report for:
Summer Faculty Research Program
Wright Laboratory Armament Directorate

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Laboratory Armament Directorate
Eglin AFB, FL

September 1996

EFFECTS OF AIRBLAST CHARACTERISTICS ON STRUCTURAL RESPONSE

Joseph W. Tedesco
Gottlieb Professor
Department of Civil Engineering
Auburn University

Abstract

The use of smaller munitions by Air Force aircraft for attacks on fixed hard target threats is receiving serious consideration. Presumably, the motivation for this action is to increase aircraft loadout, thus permitting a greater number of targets that can be destroyed in a single sortie. A major concern with the smaller munitions is their effectiveness in delivering a warhead having a blast impulse capable of destroying the majority of the fixed hard target threats. This report summarizes the results of a preliminary analytical study to assess the defining characteristics of an airblast on the structural response of a predefined hard target. The results of the investigation indicates that the structural response of the target to a specific impulse is significantly influenced by the airblast peak pressure, duration of the positive pressure phase and shape of the pressure-time curve.

EFFECTS OF AIRBLAST CHARACTERISTICS ON STRUCTURAL RESPONSE

Joseph W. Tedesco

Introduction

The effectiveness of many Air Force aircraft deployed for attacks on hard fixed targets is severely restricted by the limited number of standard sized conventional ordinance the aircraft can deliver in a single sortie. The use of smaller munitions (i.e. 250 lb class) with a general purpose warhead would permit a 3 to 4 fold increase in aircraft loadout, thereby significantly increasing the number of targets that can be destroyed in a single sortie. For a given number of aircraft, use of smaller munitions will increase the tempo of the war and allow more targets to be destroyed in a shorter time span. This would presumably result in realistic potential for abbreviating the war. Moreover, the smaller munitions would provide for greater flexibility to future aircraft designs in the sizing of the weapons bays, which controls the overall size of the aircraft.

Before the small munitions potential can be fully realized, however, several key technology issues must be addressed. One important issue is the effectiveness of the smaller munitions to deliver a warhead having a blast impulse capable of destroying the majority of the fixed hard target threats. This report summarizes the results of a preliminary analytical study to assess the effects of blast wave characteristics on the structural response of a predefined hard target. The pertinent blast characteristics considered in the study are the impulse, peak pressure and duration of the positive phase of the blast. It is anticipated that the results of this study will lend some insight to optimization of the blast pressure-time characteristics (from a lethality point of view) of small munitions.

Scope of Study

The target structure investigated in the study was a 1 ft. thick reinforced concrete slab having overall dimensions of 11 ft. x 11 ft. The slab is symmetrically reinforced with No. 9 bars top and bottom, each way, with a 2 in. clear cover. The compressive strength of the concrete, f'_c , was assumed to be 6000 psi and the yield strength of the steel, f_y , was 60 ksi. Two basic load conditions were considered. Load Case I was defined by 50 lbs of TNT detonated at a 9.5 ft standoff from the center of the slab. Load Case II corresponds to the same TNT charge at a 5.5 ft standoff. The reflected peak pressure-time histories for the positive phase of the airblasts for Load Case I and Load Case II are presented in Figure 1 and 2, respectively. The peak pressure variation over the surface of the slabs are presented in Figures 3 and 4, respectively, for Load Case I and Load Case II.

To evaluate the structural response of the target structure to the defining characteristics of the blast pressure-time histories, several combinations of the pressure time curves specified for Load Case I and Load Case II were considered. In each combination, the peak impulse I_o , was kept constant and only the peak pressure, P_o , and duration of the positive phase, t_d , were varied. The

pertinent pressure and time parameters for all combinations of each load case are summarized in Table 1. In all cases, a linearly decaying pressure-time relationship was assumed. The arrival time of the blast wave at each location on the slab was assumed to be instantaneous.

Table 1. Pressure-Time Combinatons for Load Cases I and II

LOAD CASE COMBINATION	Load Case I ($I_o = 225$ psi - msec)		Load Case II ($I_o = 460$ psi - msec)	
	Po (psi)	t_d (msec)	Po (psi)	t_d (msec)
A	600	0.75	2889.0	0.32
B	450	1.0	2166.75	0.48
C	300	1.5	1444.5	0.64
D	150	3.0	722.25	1.28

Methodology

The numerical simulation of the response of the reinforced concrete target structure to the prescribed load cases (defined in the previous section of this report) was accomplished by the finite element method (FEM) of analyses, through implementation of the ADINA[1] finite element computer program. To expedite the analyses, a typical one foot wide strip of the slab was incorporated into the FEM model. Symmetry of the structure and load cases permitted a half-span representation of the slab. The concrete was modeled with 154 two-dimensional, 8-node isoparametric plane stress elements. The reinforcing steel was represented with 3-node truss elements. A typical mesh for the FEM model is illustrated in Figure 5.

To establish the dynamic characteristics of the target structure a frequency analysis was conducted first. Next, to gain insight to the overall structural response of the target structure, linear transient analyses for Load Cases IA and IIA were conducted. Finally, material nonlinear only (MNO) transient analyses were conducted for all the Load Case combinations specified in Table 1. A summary of the FEM analyses conducted for the study are presented in Table 2.

Table 2. Summary of FEM Analyses

TYPE OF ANALYSIS	LOAD CASE							
	IA	IB	IC	ID	IIA	IIB	IIC	IID
Frequency Analysis	X							
Linear Transient Analysis	X				X			
MNO Transient Analysis	X	X	X	X	X	X	X	X

Results of Analyses

The results of the frequency analysis for the first four modes of vibration are summarized in Table 3. The mode shapes for the first 3 modes are illustrated in Figure 6. The natural period for the fundamental mode is approximately 10 msec (refer to Table 3), whereas the shortest load duration, t_d , is 0.75 msec for Load Case IA and 0.32 msec for Load Case IIA. (Refer to Table 1). This implies that the deciding factor for time step size in the numerical integration of the transient response is load dominated rather than structure dominated.

Table 3. Results of Frequency Analysis

Mode	Natural Frequency (hz)	Natural Period (msec)
1	95.41	10.48
2	451.5	2.215
3	959.9	1.042
4	1061.0	0.942

The maximum responses for several important parameters are summarized in Table 4 for the Load Case IA and IIA linear transient analyses. Time histories for midspan deflection and midspan reinforcing bar tensile stress (bottom bars), respectively, are presented in Figures 7 and 8 for Load Case IA, and in Figures 9 and 10 for Load Case IIA. The results of the linear transient analyses indicate that Load Case I does not possess sufficient impulse to impart significant structural damage to the target structure, however Load Case II appears to be catastrophic.

Table 4. Maximum Responses from Linear Transient Analyses

Maximum Response	Load Case IA	Load Case IIA
Midspan Deflection (in)	0.162	0.283
Stress in Bottom Reinforcement (ksi)	26.8	49.7
Stress in top reinforcement (ksi)	28.8	49.1
Compressive Stress in Concrete (psi)	4200	7405
Tensile Stress in Concrete (psi)	3500	6934

In an attempt to quantify the structural damage sustained by the target structure, a series of nonlinear transient analyses were conducted for each basic load case. The type of nonlinear analyses conducted is termed an MNO (Materially nonlinear only). The concrete material model prescribed in the analyses is a hypoelastic model based on a uniaxial stress-strain relation that is generalized to take biaxial and triaxial conditions into account {2, 3}. The model employs three basic features to describe the material behavior: 1) a nonlinear stress-strain relation including strain softening to allow for weakening of the material under increasing compressive stress; 2) a failure envelope that defines cracking in tension and crushing in compression; and 3) a strategy to model post-cracking and crushing behavior of the material. The parameters defining the concrete uniaxial stress-strain relation are summarized in Table 5. The material model used to describe the reinforcing steel is a bilinear elastoplastic stress-strain relation defined by an elastic modulus E , a yield stress f_y and a tangent modulus E_t . The defining parameters for the bilinear elastoplastic material model are summarized in Table 6.

For both the linear and nonlinear transient analyses solutions a direct numerical integration procedure was utilized. In the present study, the Newmark method of time integration with a consistent mass formulation was employed. The time step for temporal integration, which was controlled by the duration of the positive phase of the blast pulse, used for all analyses was 50×10^{-6} sec.

Table 5. Concrete Material Model Parameters

Parameter	Value
Initial tangent modulus, E_0 (ksi)	4.5×10^3
Uniaxial Cut-Off tensile strength, f_t (psi)	600
Uniaxial maximum compressive strength, f'_c (psi)	6000
Uniaxial ultimate compressive strength, f_u (psi)	5100
Uniaxial compressive strain at f'_c	0.002
Uniaxial compressive strain at f_u	0.003

Table 6. Reinforcing Steel Material Model Parameters

Parameter	Value
Initial elastic modulus, E (ksi)	29×10^3
Tangent modulus, E_t (ksi)	300
Yield Strength, f_y (ksi)	60
Yield strain	0.00207
Poisson's ratio	0.3

Because of the complexity of the material descriptions used in the MNO transient analyses, an appropriate strategy for solving the nonlinear analyses, the specifically full Newton-Raphson

interaction scheme with line searches [4, 5], was employed. To minimize the number of stiffness matrix reformulations, the automatic time stepping (ATS) option was activated with a maximum of 50 subincrements of each time step.

The maximum responses for midspan deflection, tensile stress in the reinforcement (bottom bars) at midspan, tensile stress in the reinforcement (top bars) at the support and the time to failure, obtained from the nonlinear transient analyses, are summarized in Table 7. The time histories for the midspan deflection for Load Cases IA and ID are presented in Figure 11, and those for Load Cases IIA and IID are shown in Figure 12. The end of simulation crack patterns for Load Cases IA through ID are presented in Figure 13, and the crack patterns at time of failure for Load Cases IIA through IID are shown in Figure 14.

Table 7. Summary of Nonlinear Analyses Results

Load Case	Midspan deflection (in)	Stress in rebar at midspan (ksi)	Stress in rebar at support (ksi)	Time to failure (msec)
IA	0.246	28.11	38.63	-
IB	0.229	30.2	39.61	-
IC	0.242	26.34	50.17	-
ID	0.241	34.65	55.05	-
IIA	0.0387	2.48	32.7	0.364
IIB	0.108	9.68	31.89	0.774
IIC	0.122	5.5	41.72	1.032
IID	0.405	59.84	60.14	2.975

Discussion of Results

Several interesting observations can be made from the results of the nonlinear transient analyses. For the low impulse load case, Load Case I, it appears that even though the loading condition does not fail the structure, the extent of structural damage is increased by increasing the time over which the load acts on the structure. However, the optimal pressure-time combination for maximum structural damage for a specified impulse is not readily apparent from these limited

number of case studies. The results from the high impulse load case, Load Case II, exhibit very different trends. Damage appears to be complete and catastrophic when the impulse is delivered with high intensity and short duration. As illustrated in Figure 14a and Table 7, most of the concrete in the structure has cracked or crushed, and very low stress levels have developed in the reinforcement. This type of response indicates a complete "shear out" failure condition. When the impulse is delivered over a longer time span, more structural response is experienced by the slab, as indicated in Figure 14b and Table 7. The reinforcement appears to have yielded at the support and midspan, accompanied by significant cracking and crushing at these locations. This behavior indicates a hinge mechanism failure, with some probable localized breach.

Conclusions and Recommendations

The results of this study clearly indicate the response of structures to impulse loading is strongly influenced by the manner in which the impulse is delivered; that is, the relative magnitudes of the peak pressure and time of duration. For high impulse loads, damage is maximized by delivering the impulse over a short time of duration. Extending the time over which the impulse is delivered appears to invoke more structural response and transition the failure mode from a complete, catastrophic shear-out failure toward a flexural yield line failure accompanied by some localized breaching. In the case of a low impulse load, one in which a shear-out failure or breach is not instigated, extending the time over which the impulse is delivered appears to maximize the structural damage. However, the optimum peak pressure and impulse duration cannot be fully realized from the results of this study. Moreover, it is the author's opinion that the shape of the pressure-time curve may also have a significant effect on structural response, especially for longer duration impulses.

It is recommended that the scope of this study be expanded to include a larger number of different types of impulsive loads and target structures. Future studies should also direct attention to lower impulse loads in an enclosed structure, where more complex impulsive load histories are generated. Finally, to more accurately simulate the structural response of the actual structure, three dimensional simulations should be conducted in future studies.

Acknowledgments

This research was sponsored by the Air Force Office of Scientific Research, Bolling Air Force Base, D.C. Computational resources for the numerical simulations were provided by Wright Laboratory Armament Directorate, WL/MNSA, Eglin AFB, FL.

References

1. ADINA, A Finite Element Computer Program for Automatic Dynamic Incremental Nonlinear Analysis, Report ARD 90-2, ADINA R&D, Inc, Watertown, MA, 1987.
2. Bathe, K. J., Walezak, J. Welch, A. And Mistry, N., "Nonlinear Analysis of Concrete Structures", Journal of Computers and Structures, Vol. 32, No. 3, pp 563-590, 1989.
3. Bathe, K.J. and Ramaswamy, S., "On Three Dimensional Nonlinear Analysis of Concrete Structures", Nuclear Engineering and Design, Vol. 52, pp 385-409, 1979.
4. ADINA Theory and Modeling Guide, Report ARD 87-8, ADINA R&D, Watertown, MA, 1987.
5. Bathe, K.J., Finite Element Procedures in Engineering Analysis, Prentice-Hall Inc., Englewood Cliffs, NJ, 1982.

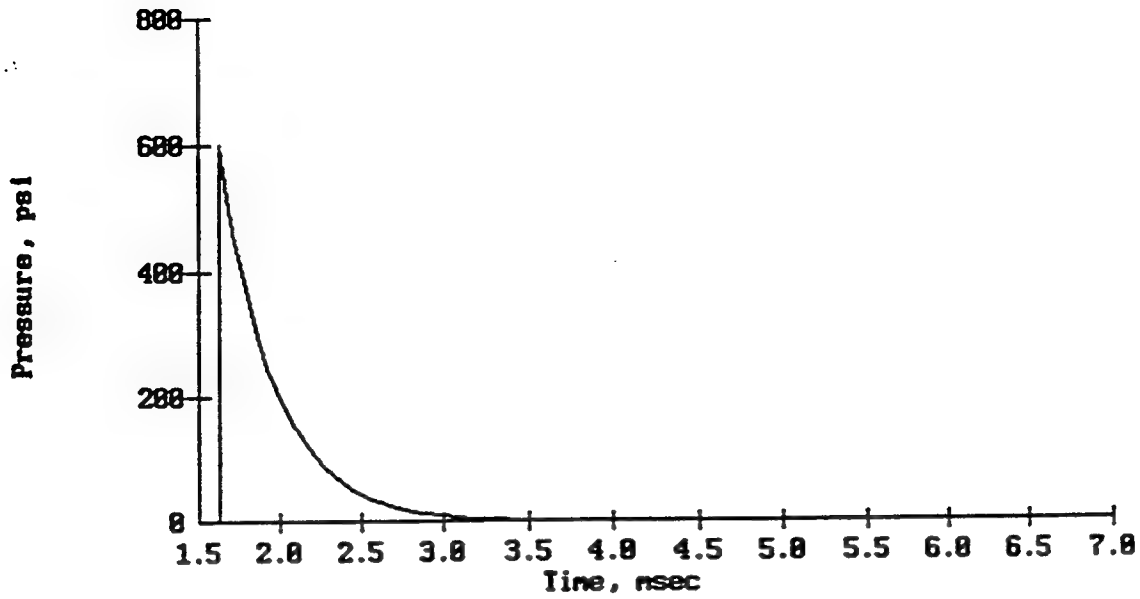


Figure 1: Reflected Pressure-Time History for 50 lbs of TNT at 9.5 ft. Standoff

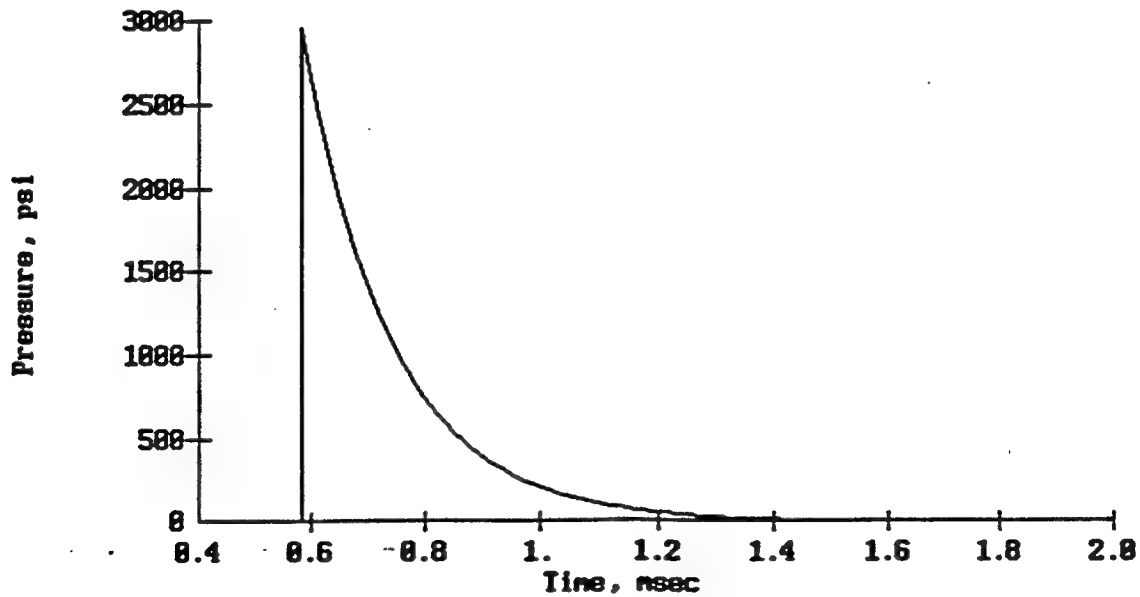


Figure 2. Reflected Pressure-Time History for 50 lbs of TNT at 5.5 ft. Standoff

Pressure vs Distance

50# TNT @ 9.5 ft standoff
11' X 11' wall

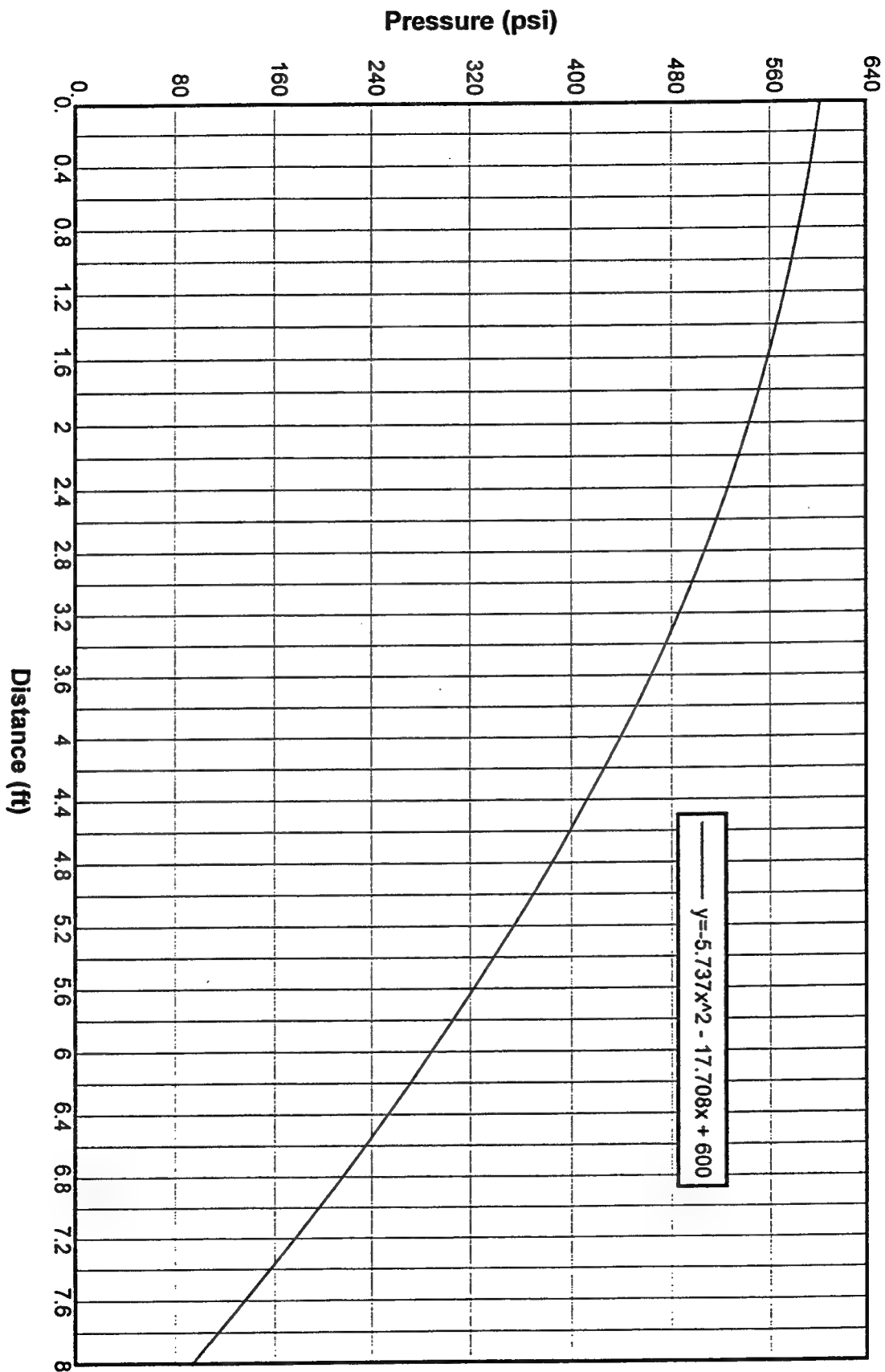


Figure 3. Peak Reflected Pressure Distribution over Target Structure, Load Case I

Pressure vs Distance **50# TNT @ 5.5 ft standoff** **11' X 11' wall**

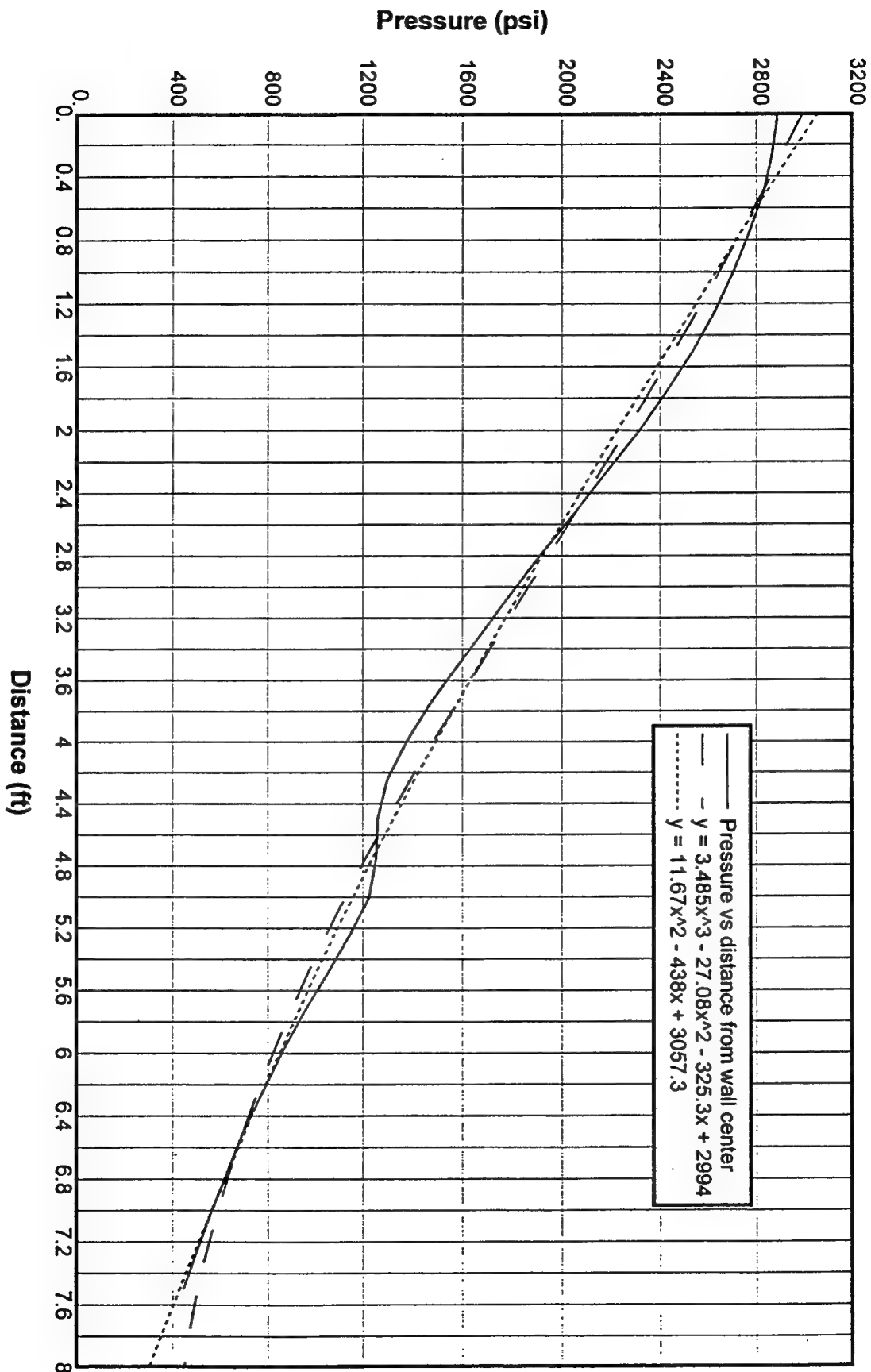


Figure 4. Peak Reflected Pressure Distribution over Target Structure, Load Case II

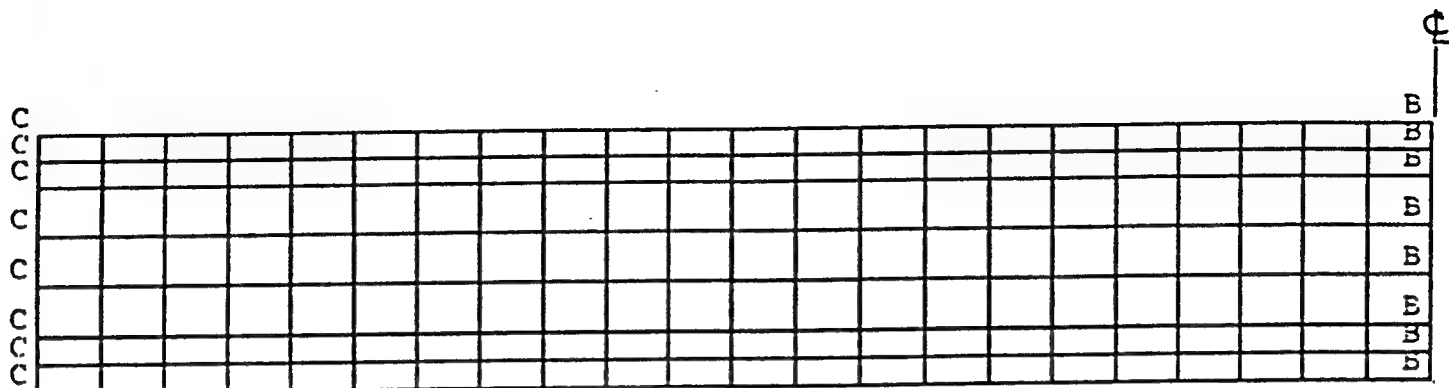


Figure 5. FEM Mesh of Target Structure

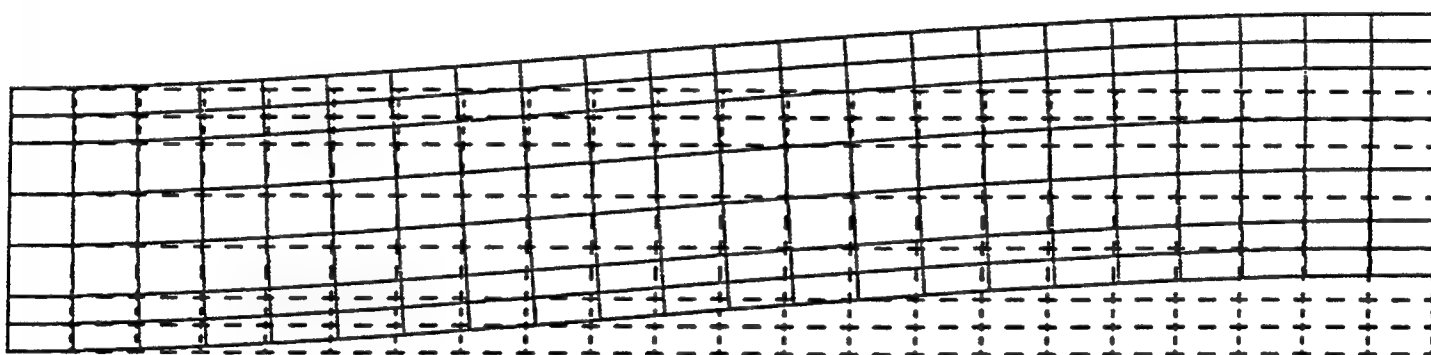


Figure 6a. Mode Shape for Mode 1.

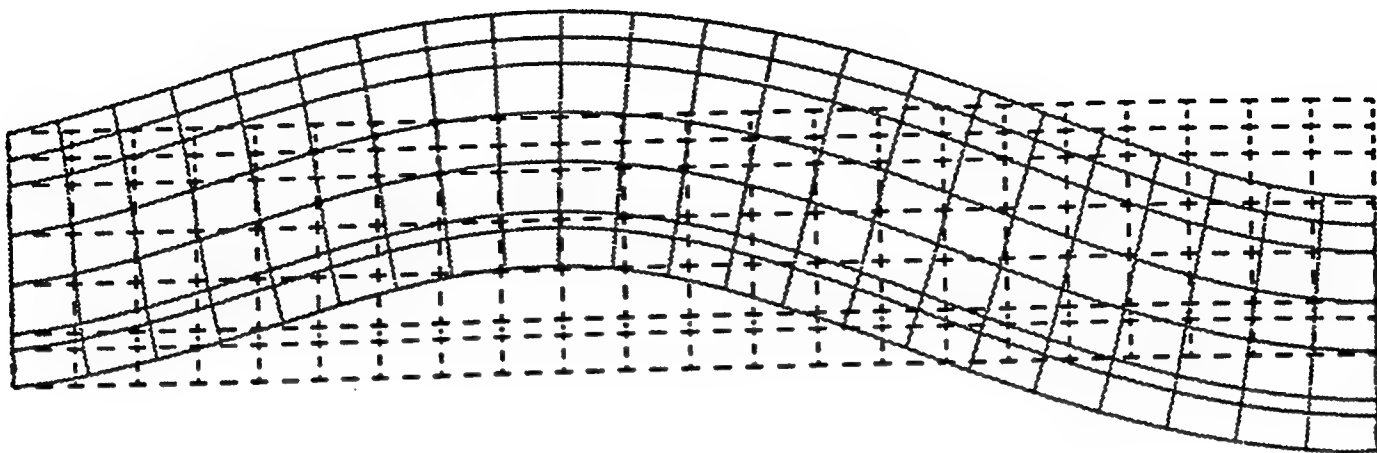


Figure 6b. Mode Shape for Mode 2

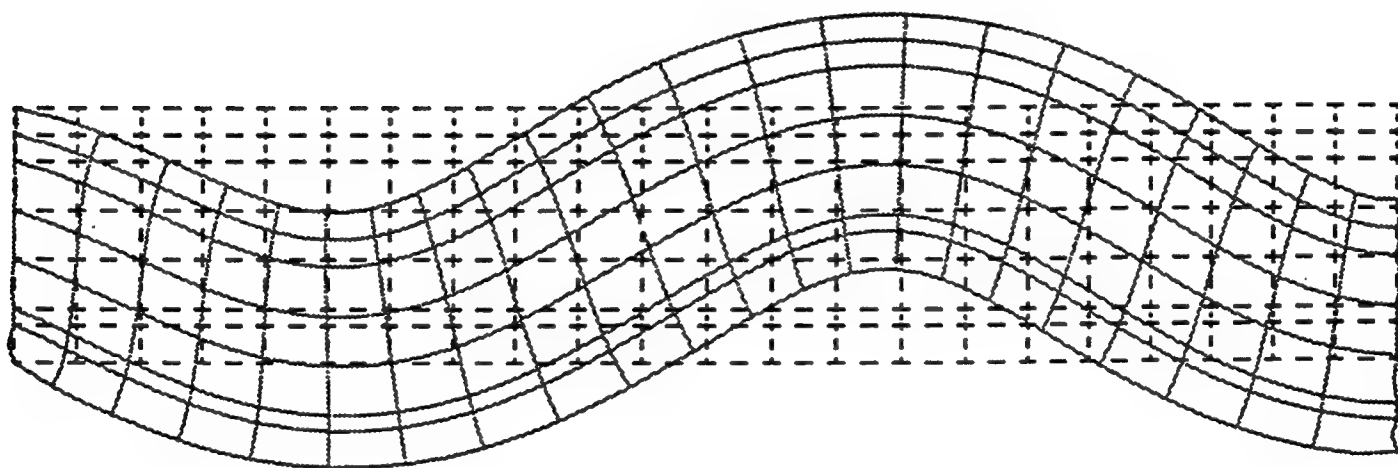


Figure 6c. Mode Shape for Mode 3

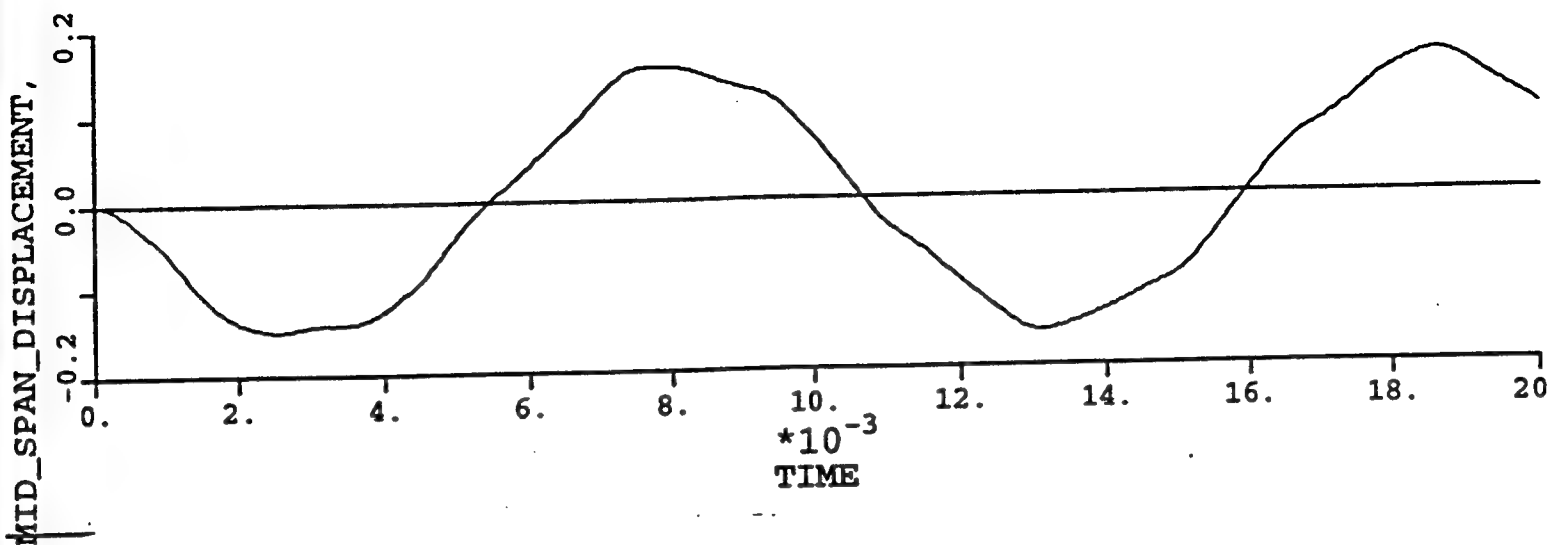


Figure 7. Midspan Deflection (in.) of Target Structure, Linear Transient Analysis;
Load Case I

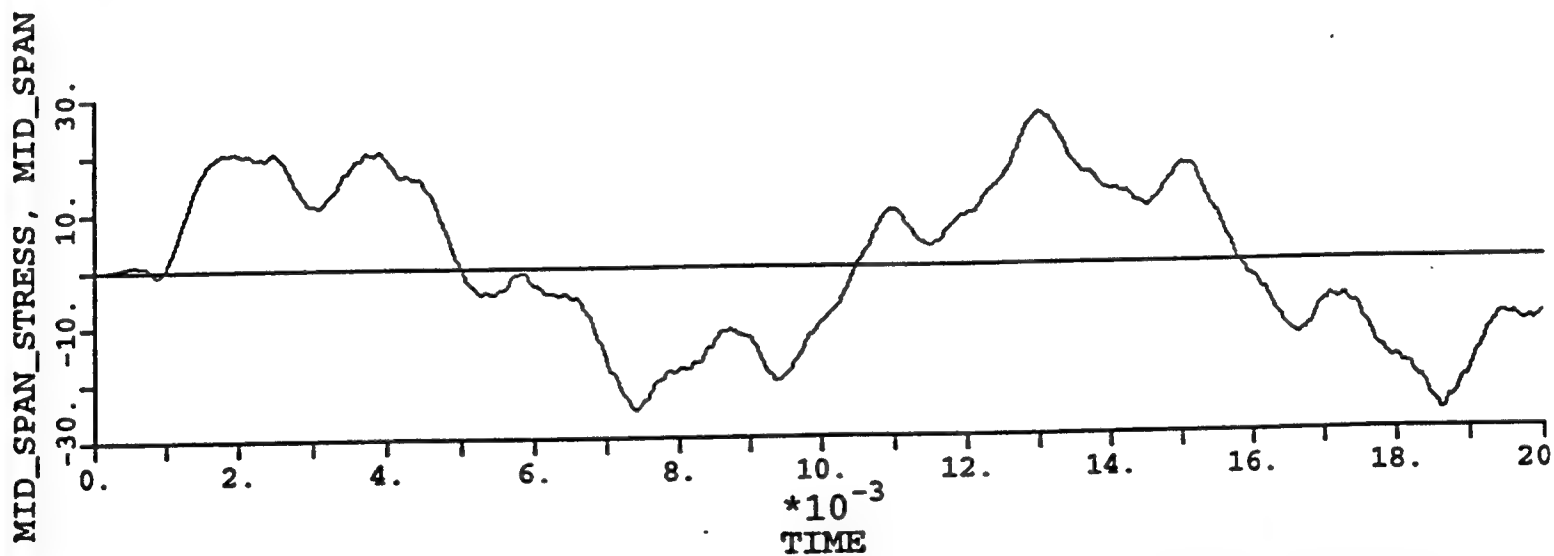


Figure 8. Stress in Reinforcing Steel (ksi) at Midspan, Linear Transient Analysis;
Load Case I

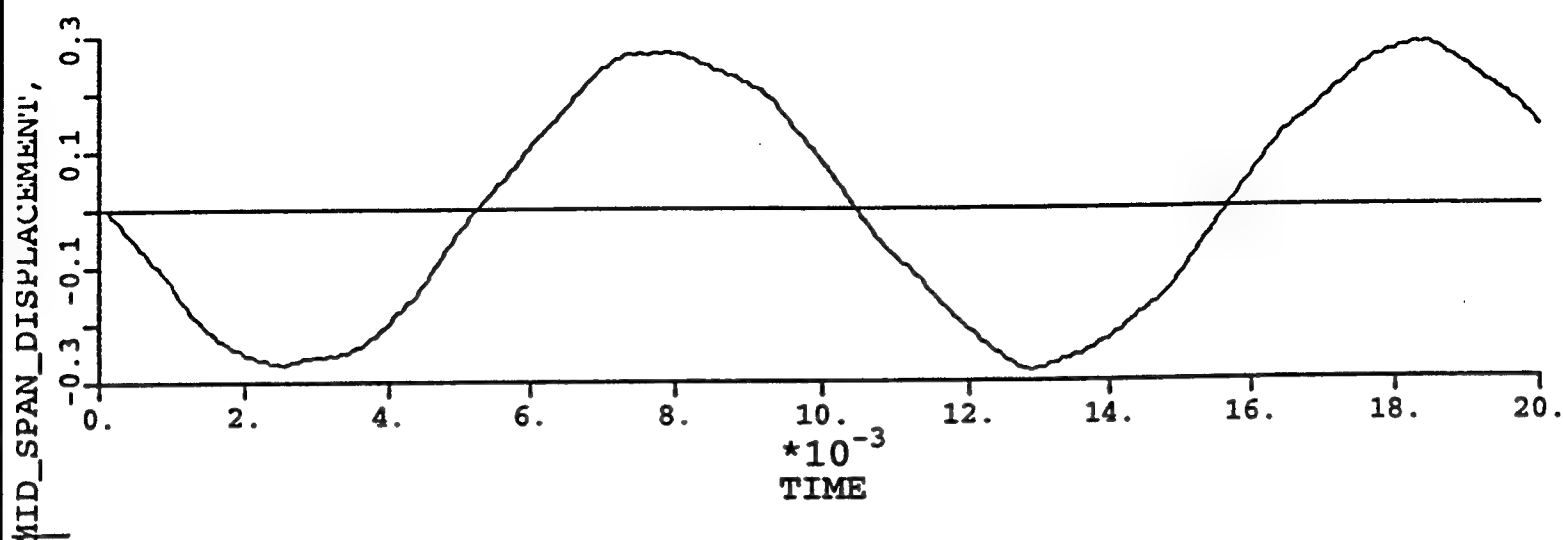


Figure 9. Midspan Deflection (in.) of Target Structure, Linear Analysis;
Load Case II

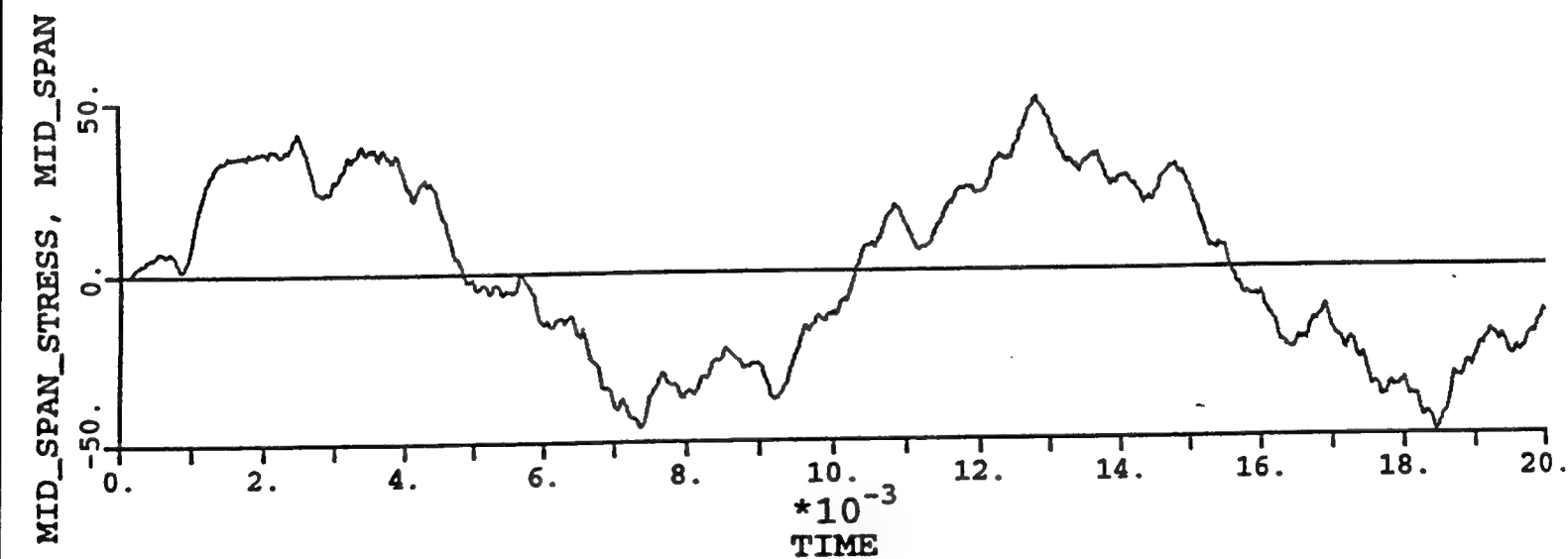


Figure 10. Stress in Reinforcing Steel (ksi) at Midspan, Linear Transient Analysis;
Load Case II

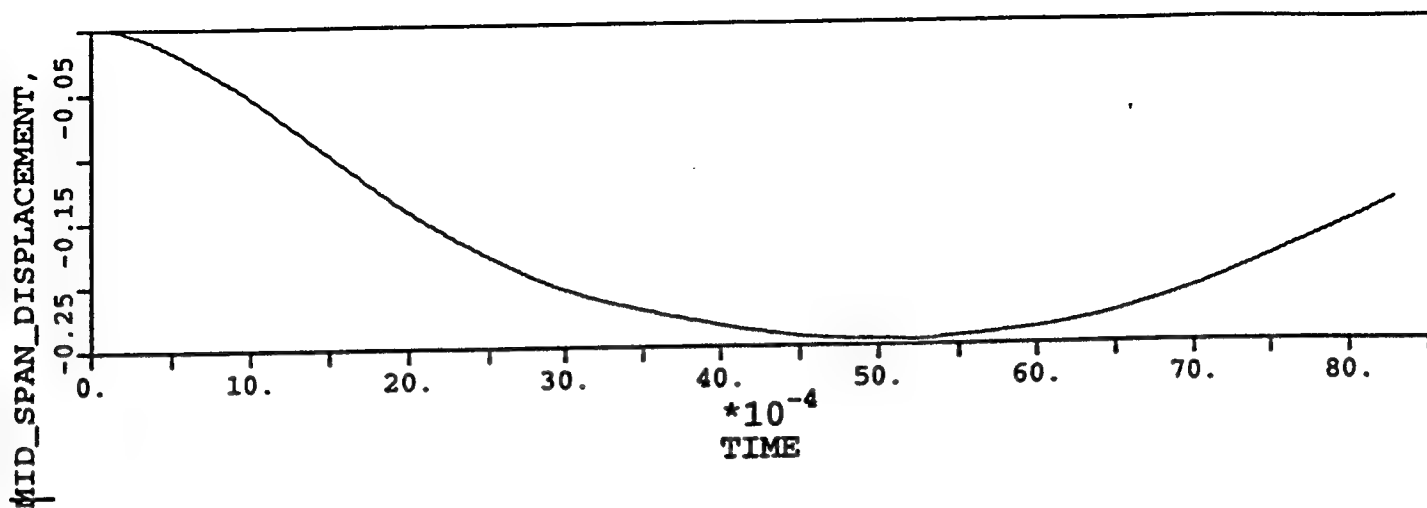


Figure 11a. Midspan Deflection (in.) of Target Structure, Nonlinear Transient Analysis; Load Case 1A

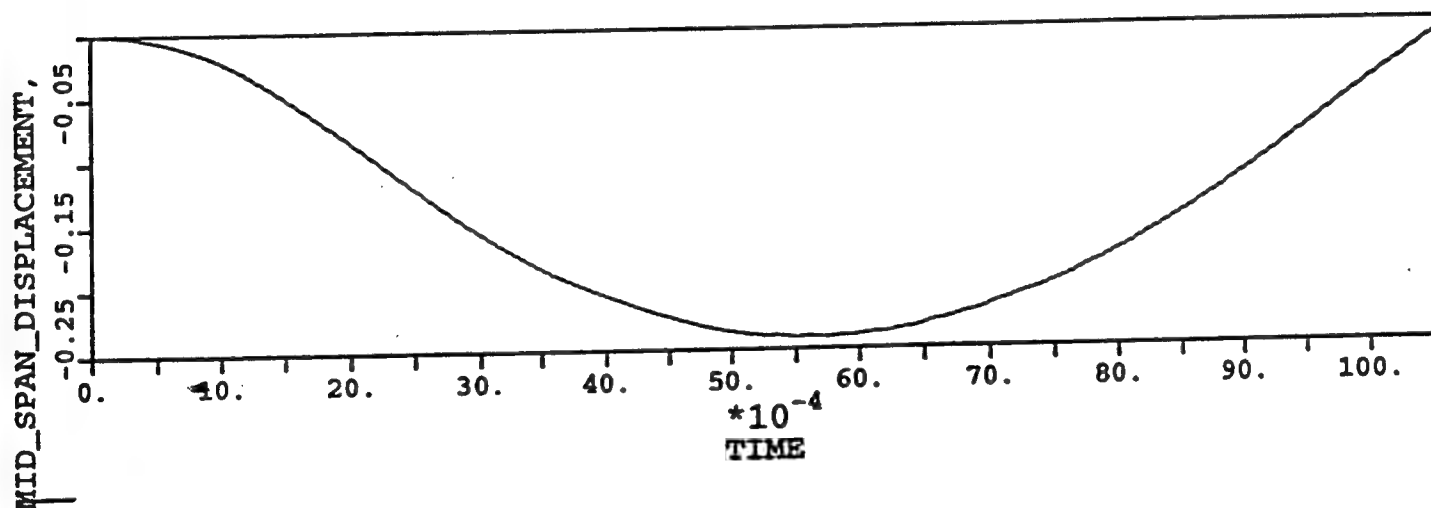


Figure 11b. Midspan Deflection (in.) of Target Structure, Nonlinear Transient Analysis, Load Case 1D

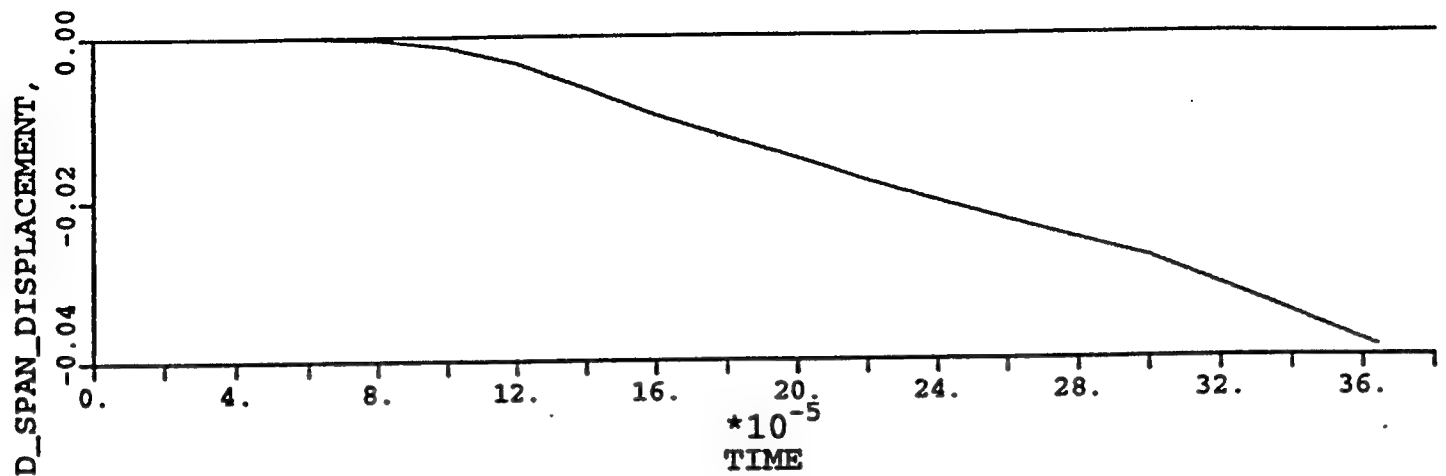


Figure 12a. Midspan Deflection (in.) of Target Structure, Nonlinear Transient Analysis;
Load Case IIA

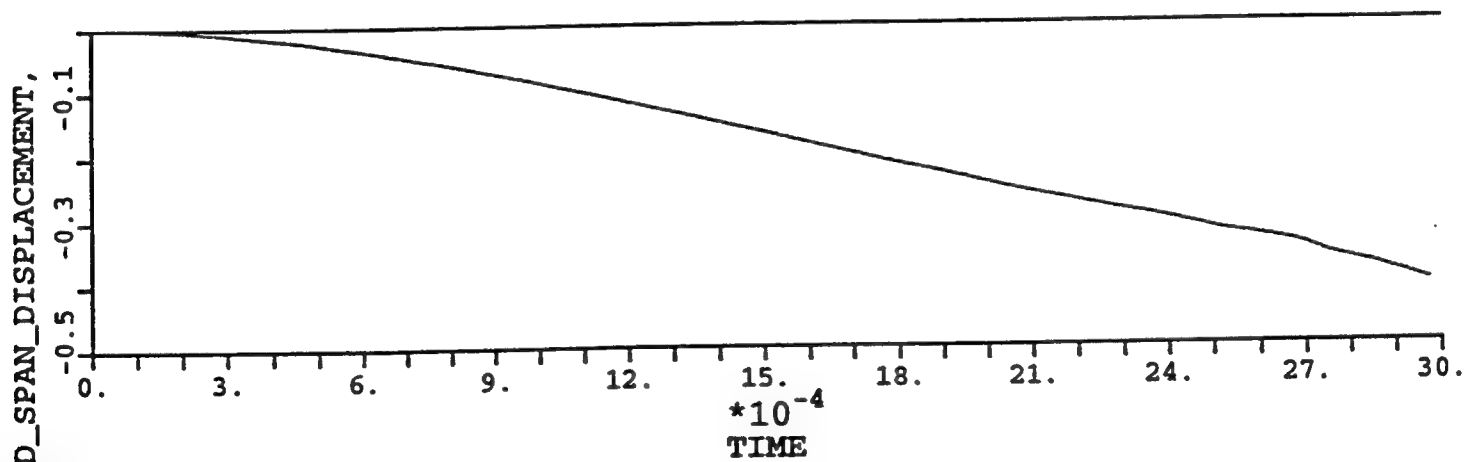
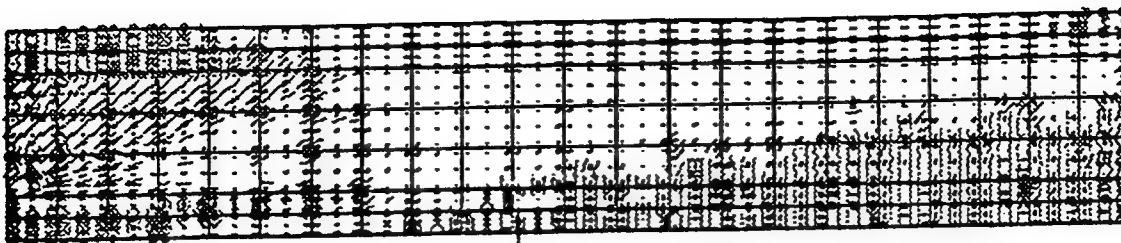
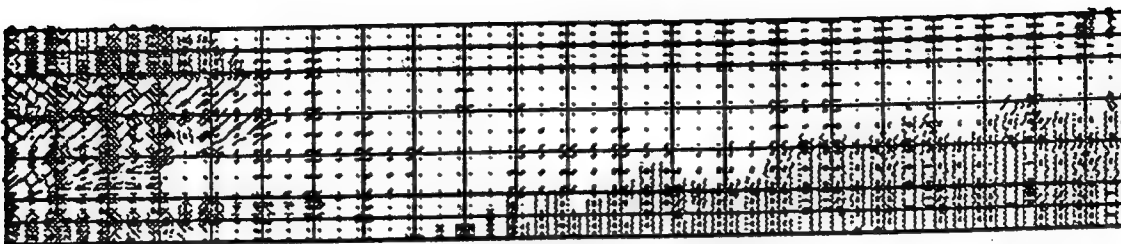


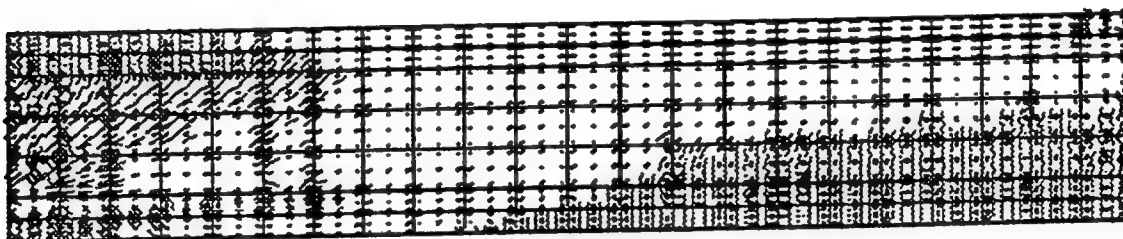
Figure 12b. Midspan Deflection (in.) of Target Structure, Nonlinear Transient Analysis;
Load Case IID



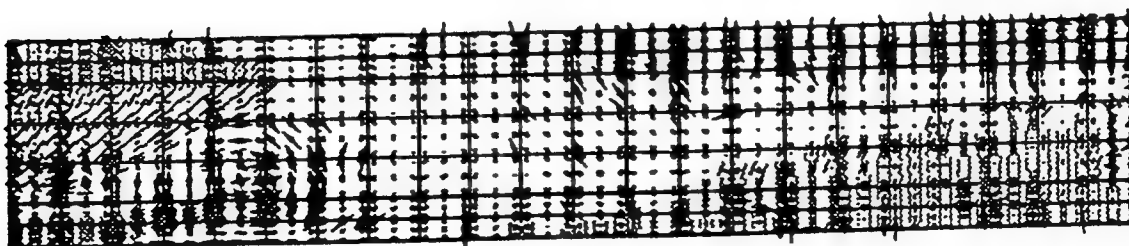
a) IA



b) IB



c) IC



d) ID

Figure 13. End of Simulation Crack Patterns for Load Case I,
Nonlinear Transient Analysis

THE EFFECTS OF CURVATURE ON THE PERFORMANCE OF A
SPIRALLY-GROOVED COPPER-ETHANOL HEAT PIPE

Scott K. Thomas
Assistant Professor
Department of Mechanical and Materials Engineering

Wright State University
Dayton, OH 45435
Ph: (513) 873-5040
Fax: (513) 873-5009

Final Report for:
Summer Faculty Research Program
Wright Laboratories

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright-Patterson Air Force Base, Dayton OH

August 1996

THE EFFECTS OF CURVATURE ON THE PERFORMANCE OF A SPIRALLY-GROOVED COPPER-ETHANOL HEAT PIPE

Scott K. Thomas
Assistant Professor
Department of Mechanical and Materials Engineering
Wright State University

Abstract

Heat pipes are simple, effective devices which are capable of transporting heat energy from one location to another with a very low superheat penalty. Heat pipes have become excellent candidates for passive cooling of high performance aircraft. Due to high transient acceleration fields present in such applications, heat pipes must be able to operate effectively under adverse conditions. In an effort to understand heat pipe performance under transient acceleration fields, a two-phase study is being conducted. Phase 1 included the design, fabrication and testing of a spirally-grooved copper heat pipe with ethanol as the working fluid. The heat pipe was tested horizontally for the capillary limit before and after being bent to a radius of 1.22 m along the centerline of the pipe, which provided information into the effects of curvature on the heat pipe performance. During Phase 2, the transient response of the curved heat pipe will be examined on a centrifuge table with a 1.22 m radius. The centrifuge table will be able to vary the radial and tangential velocities in a sinusoidal manner as a function of time with varying frequency. By operating the centrifuge table in this manner the acceleration fields produced by high performance aircraft can be approximated. This report describes the experimental procedures and results determined during Phase 1.

THE EFFECTS OF CURVATURE ON THE PERFORMANCE OF A SPIRALLY-GROOVED COPPER-ETHANOL HEAT PIPE

Scott K. Thomas

Introduction

The analysis of heat pipes has usually been restricted to the inclusion of a static acceleration field, such as that due to gravity. While this analysis is appropriate in many applications, it is not valid in the assessment of the thermal performance of heat pipes in acceleration fields which are varying with time. For instance, heat pipes have been proposed to be used aboard fighter aircraft such as the Navy F/A-18 to act as heat sinks for electronics packages which drive aileron or trailing edge flap actuators (Gernert et al., 1991; Yerkes and Beam, 1992; Yerkes and Hager, 1992). During combat, transient acceleration fields of up to 9-*g*'s are present on the aircraft. Therefore, knowledge of the thermal performance of heat pipes under transient acceleration loadings is of importance to designers of the electronics packages in need of cooling. Thomas and Yerkes (1996) determined the quasi-steady state thermal resistance of a flexible copper-water heat pipe under transient acceleration loadings with constant heat input using a horizontal centrifuge table. The performance of the heat pipe was examined for the following parameter ranges: Heat input, $Q_e = 75$ to 150 W; condenser temperature, $T = 3, 20$ and 35°C ; and acceleration frequency, $f = 0, 0.01, 0.05, 0.1, 0.15$ and 0.2 Hz. The centrifuge radial acceleration loadings ranged from 1.1 to 9.8-*g*'s for each frequency setting. In addition, the effects of the previous dryout history were noted. It was found that the thermal resistance of the heat pipe was a function of the acceleration frequency, heat input, condenser temperature, and dryout condition prior to changing the frequency. The objective of the present experimental study is to further the analysis by Thomas and Yerkes (1996) by examining the thermal characteristics of a heat pipe with a well-defined wick structure (axial grooves). The results of this experimental analysis can then be used for comparison with future analytical models of the heat pipe system under transient acceleration loadings.

Mathematical Model

The key to calculating the capillary limit of a heat pipe is understanding the pressure balance between the vapor and liquid in the heat pipe. Since the rate of twist of the spiral grooves within the present heat pipe wall was only $2.6^\circ/\text{cm}$, and the rate of curvature of the curved heat pipe was very shallow, the capillary limit was modeled for a heat pipe with a straight axially-grooved wick structure.

The pressure balance within the heat pipe can be written as

$$[P_v(x_{\text{ref}}) - P_v(x)] + [P_v(x) - P_l(x)] + [P_l(x) - P_l(x_{\text{ref}})] + [P_l(x_{\text{ref}}) - P_v(x_{\text{ref}})] = 0 \quad (1)$$

When the capillary pressure, P_c , which is defined as the pressure at the vapor side of the liquid-vapor interface minus that at the liquid side (Chi, 1976), is introduced into equation (1) with $x_{\text{ref}} = x_{\text{min}}$

$$P_c(x) = \Delta P_v(x - x_{\text{min}}) + \Delta P_l(x_{\text{min}} - x) \quad (2)$$

or

$$P_c(x) = \int_{x_{\text{min}}}^x \left(\frac{dP_v}{dx} - \frac{dP_l}{dx} \right) dx \quad (3)$$

where ΔP_v and ΔP_l are the vapor and liquid pressure changes between x and x_{min} . By examining the shape of the meniscus formed at the liquid-vapor interface in the axial and circumferential directions, the capillary pressure can be written as

$$P_c = P_v - P_l = \sigma \left(\frac{1}{R_{\text{circ}}} + \frac{1}{R_{\text{axial}}} \right) \quad (4)$$

where R_{axial} and R_{circ} are the principal radii of curvature in the axial and circumferential directions, respectively, and σ is the liquid surface tension. For an axial groove with a rectangular cross section, $R_{\text{axial}} = \infty$. The maximum capillary pressure is obtained when the meniscus formed by the liquid-vapor interface is tangent to the groove wall. Therefore, the capillary radius equals the groove width

$$r_c = w \quad (5)$$

and the maximum capillary pressure is

$$P_{\text{cm}} = \sigma \left(\frac{1}{R_{\text{circ}}} \right) = \frac{2\sigma}{r_c} \quad (6)$$

The above equation assumes no inclination of the heat pipe and no circumferential communication of the liquid. The pressure drop around the circumference of the wick structure is defined as

$$\Delta P_{\perp} = \rho_l g d_v \cos \psi \quad (7)$$

where ψ is the heat pipe inclination angle and d_v is the vapor core diameter. For the present case, $\Delta P_{\perp} = 0$ because of the axially-grooved wick structure. With the effects of the pressure drop around the circumference of the wick structure taken into account, equation (6) becomes

$$P_{\text{cm}} = \frac{2\sigma}{r_c} - \Delta P_{\perp} = \int_0^{L_t} \left(\frac{dP_v}{dx} - \frac{dP_l}{dx} \right) dx \quad (8)$$

with $x_{\min} = 0$ and $x = L_t$. Therefore, the general expression for the maximum capillary pressure is

$$P_{\text{cm}} = \frac{2\sigma}{r_c} - \Delta P_{\perp} - \Delta P_v - \Delta P_l - \rho_l g L_t \sin \psi \quad (9)$$

In order to determine the capillary limit, the pressure gradients in the vapor and liquid, dP_v/dx and dP_l/dx , must be found. The vapor pressure gradient is modeled by

$$\Delta P_v(x - x_{\min}) = P_v(x) - P_v(x_{\min}) = \int_{x_{\min}}^x \frac{dP_v}{dx} dx \quad (10)$$

Faghri and Parvani (1987) stated that fully-developed vapor flow is obtained after a very short distance from the evaporator end cap. In this numerical study, Faghri and Parvani developed equations to model the pressure drop in all three sections of a conventional heat pipe:

$$P_{\text{evap}} = E_1 x_{\text{evap}}^2 + P_{\text{in, evap}} \quad (11)$$

$$P_{\text{adiab}} = A_1 x_{\text{adiab}} + P_{\text{in, adiab}} \quad (12)$$

$$P_{\text{cond}} = C_1 x_{\text{cond}}^2 + C_2 x_{\text{cond}} + P_{\text{in, cond}} \quad (13)$$

where $0 \leq x_{\text{evap}} \leq L_{\text{evap}}$, $0 \leq x_{\text{adiab}} \leq L_{\text{adiab}}$, and $0 \leq x_{\text{cond}} \leq L_{\text{cond}}$. The inlet pressures for the three sections of the heat pipe are $P_{\text{in, evap}}$, $P_{\text{in, adiab}}$, and $P_{\text{in, cond}}$. The coefficients E_1 , A_1 , C_1 , and C_2 are defined by

$$E_1 = \left(-8\text{Re}_{\text{evap}} - \frac{16}{3}\text{Re}_{\text{evap}}^2 \right) \left(\frac{\mu_v^2}{\rho_v r_{h,v}^4} \right) \quad (14)$$

$$A_1 = \left(\frac{-8\mu_v \bar{w}_a}{r_{h,v}^2} \right) \quad (15)$$

$$C_1 = \left(8\text{Re}_{\text{cond}} - \frac{16}{3}\text{Re}_{\text{cond}}^2 \right) \left(\frac{\mu_v^2}{\rho_v r_{h,v}^4} \right) \quad (16)$$

$$C_2 = \left(\frac{16}{3}\text{Re}_{\text{cond}}^2 - 8 \right) \left(\frac{-8\mu_v \bar{w}_a}{r_{h,v}^2} \right) \quad (17)$$

where the hydraulic radius of the vapor core is $r_{h,v} = d_v/2$ and the Reynolds numbers, Re_{evap} and Re_{cond} , as well as the axial mean velocity in the adiabatic section, \bar{w}_a , can be written as

$$\text{Re}_{\text{evap}} = \frac{\rho_v U_{\text{evap}} r_{h,v}}{\mu_v} = \frac{Q}{2\pi L_{\text{evap}} h_{fg} \mu_v} \quad (18)$$

$$\text{Re}_{\text{cond}} = \frac{\rho_v U_{\text{cond}} r_{h,v}}{\mu_v} = \frac{Q}{2\pi L_{\text{cond}} h_{fg} \mu_v} \quad (19)$$

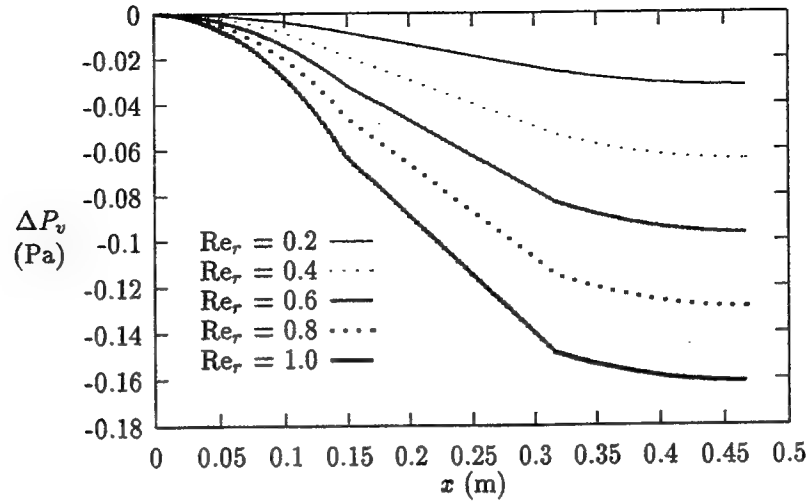


Figure 1: Pressure drop versus axial length for various radial Reynolds numbers. The working fluid is water and $Re_r = Re_{evap} = Re_{cond}$.

$$\bar{w}_a = \frac{Q}{\rho_v \pi r_{h,v}^2 h_{fg}} \quad (20)$$

U_{evap} and U_{cond} are radial vapor velocities in the evaporator and condenser sections. The pressure distributions presented in Figure 1 were obtained using equations (11)–(13). The evaporator, transport section, and condenser lengths were 0.15, 0.1673, and 0.15 m, respectively. The vapor core diameter was 0.01166 m and the properties of water were established at 100°C. Notice that the calculations were made from $x = 0$ to the respective total length of each section. The three pressure distributions were then plotted at the correct location in the domain. Using equations (11)–(13), the pressure drop in each section can be calculated directly.

$$\Delta P_{evap} = E_1 L_{evap}^2 \quad (21)$$

$$\Delta P_{adiab} = A_1 L_{adiab} \quad (22)$$

$$\Delta P_{cond} = C_1 L_{cond}^2 + C_2 L_{cond} \quad (23)$$

The overall vapor pressure drop is

$$\Delta P_v = \Delta P_{evap} + \Delta P_{adiab} + \Delta P_{cond} = \frac{8Q\mu_v}{\pi\rho_v h_{fg} r_{h,v}^4} (0.5L_{evap} + L_{adiab} + 0.5L_{cond}) \quad (24)$$

Therefore, the axial pressure gradient is given by

$$\frac{dP_v}{dx} = \frac{8\mu_v}{\pi\rho_v h_{fg} r_{h,v}^4} Q = F_v Q \quad (25)$$

where the vapor friction coefficient is

$$F_v = \frac{8\mu_v}{\pi\rho_v h_{fg} r_{h,v}^4} \quad (26)$$

The liquid pressure gradient can be modeled by

$$\Delta P_l(x_{\min} - x) = P_l(x_{\min}) - P_l(x) = - \int_{x_{\min}}^x \frac{dP_l}{dx} dx \quad (27)$$

Since the liquid velocity is very low, the dynamic pressure is neglected (Chi, 1976). Therefore, the pressure gradient in the liquid is

$$\frac{dP_l}{dx} = - \frac{(f_l Re_l) \mu_l}{2\epsilon A_w r_{h,l}^2 \rho_l h_{fg}} Q \pm \rho_l g \sin \psi \quad (28)$$

$$= -F_l Q \pm \rho_l g \sin \psi \quad (29)$$

where

$$F_l = \frac{\mu_l}{K A_w h_{fg} \rho_l} \quad (30)$$

The permeability is defined by Faghri (1994) as

$$K = \frac{2\epsilon r_{h,l}^2}{(f_l Re_l)} \quad (31)$$

where the liquid hydraulic radius is

$$r_{h,l} = \frac{2\delta w}{(2\delta + w)} \quad (32)$$

The porosity in general terms is defined as the ratio of pore volume, V_p , to the total wick volume, V_t (Faghri, 1994).

$$\epsilon = \frac{V_p}{V_t} \quad (33)$$

For an axial groove with a rectangular cross section, $V_p = w\delta L_t n$ and $V_t = A_w L_t$. Therefore, porosity can be written as

$$\epsilon = \frac{w\delta L_t n}{A_w L_t} = \frac{w\delta n}{A_w} \quad (34)$$

where cross-sectional area of the axially grooved wick in a heat pipe with a circular cross section is

$$A_w = \frac{\pi}{4} [(d_v + 2\delta)^2 - d_v^2] \quad (35)$$

The coefficient of drag is given by Shah and Bhatti (1987)

$$f_l Re_l = 24(1 - 1.3553\alpha + 1.9467\alpha^2 - 1.7012\alpha^3 + 0.9564\alpha^4 - 0.2537\alpha^5) \quad (36)$$

where the aspect ratio is, $\alpha = w/\delta$ ($0 \leq \alpha \leq 1$). This polynomial curve fit is accurate to within 0.05%.

Using equation (8) and substituting equations (25) and (29) for dP_v/dx and dP_l/dx , respectively, gives the expression for the capillary limit

$$P_{cm} = \frac{2\sigma}{r_c} - \Delta P_L = \int_0^{L_t} (F_v Q + F_l Q + \rho_l g \sin \psi) dx \quad (37)$$

or

$$\frac{2\sigma}{r_c} - \Delta P_L - \rho_l g L_t \sin \psi = \int_0^{L_t} (F_v + F_l) Q dx \quad (38)$$

In terms of the heat transport factor, the capillary limitation can be defined as (Chi, 1976)

$$(QL)_{cm} = \int_0^{L_t} Q dx = \frac{2\sigma/r_c - \Delta P_L - \rho_l g L_t \sin \psi}{F_v + F_l} \quad (39)$$

where F_v and F_l are defined by equations (26) and (30), respectively. Therefore, the capillary limit is

$$Q_{cm} = \frac{(QL)_{cm}}{0.5L_{cond} + L_{adiab} + 0.5L_{evap}} \quad (40)$$

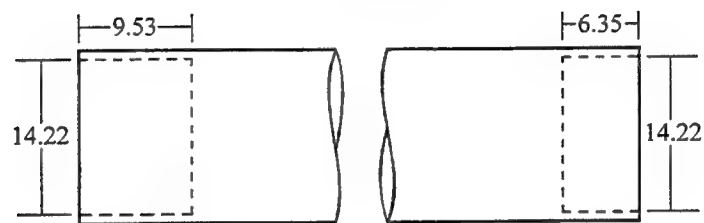
Experimental Setup

The purpose of this experiment was to establish data on the capillary limit and thermal resistance of a copper-ethanol heat pipe with spiral grooves. In order to understand the effects of curvature, the pipe had to first be studied in the straight configuration. The heat pipe was then bent to a radius of curvature of 1.22 m to match that of the WL/POOS centrifuge table. This will ensure uniform acceleration fields throughout the heat pipe during dynamic testing.

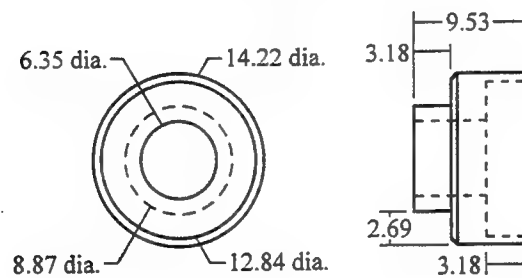
Physical information about the heat pipe is given in Table 1. Figure 2 shows the detailed drawings of the heat pipe container and end caps. To measure the performance of the heat pipe, several pieces of instrumentation were used. Included was a calorimeter, thermocouples, and power analyzers. A schematic of the setup is shown in Figure 3. The calorimeter consisted of a length of 3/16 in. copper tubing wound tightly around the condenser section of the heat pipe in which Type T thermocouples were inserted through brass T-branch connectors into the cooling water inlet and exit stream along with a flowmeter to measure the coolant flow rate. To ensure accuracy within the equipment, the flowmeter was calibrated at three temperatures (5, 15, and 35°C). The calibration procedure consisted of weighing the amount of water collected in one minute at a particular flow setting. The calibration was repeated for a total of five iterations at each flow setting and temperature. The temperature of the coolant was maintained to within $\pm 0.1^\circ\text{C}$ using the recirculating chiller.

Type T surface thermocouples were used along the pipe except on the end caps where Type T beaded thermocouples were potted in with thermal adhesive. The surface thermocouples were mounted to the heat pipe using high temperature Kapton tape and the relative positions of the thermocouples are given in

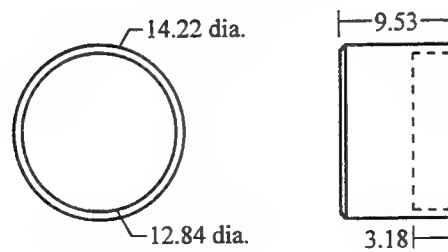
Heat Pipe Container



Fill Valve End Cap



Plain End Cap



All Dimensions in mm

Figure 2: Heat pipe container and end cap drawings.

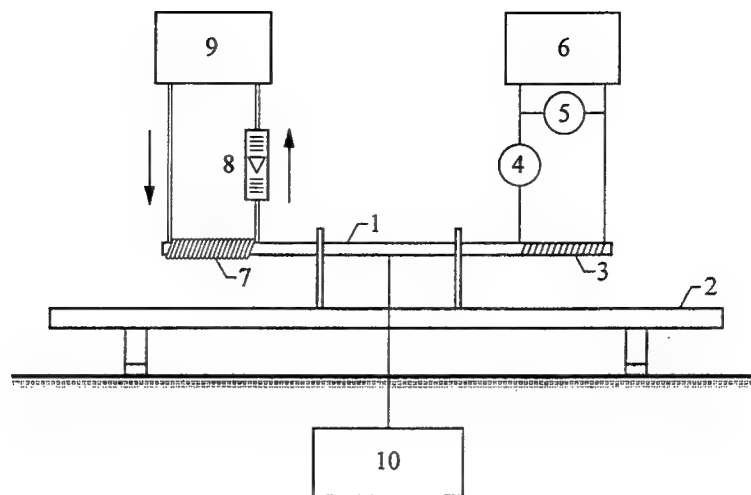
Table 1: Spirally-grooved heat pipe specifications.

Heat pipe length	450.8 mm
Evaporator length	152.4 mm
Adiabatic length	146.0 mm
Condenser length	152.4 mm
Working fluid	Ethanol
Working fluid charge	6.5 cm ³
Wall/wick materials	Copper
Wick structure	Spiral grooves
Groove twist rate	2.60 arc degrees per cm
Groove depth	0.442 mm
Groove width	0.4735 mm
Number of Grooves	50
Tube outside diameter	15.88 mm
Vapor core diameter	13.46 mm
Tube wall thickness	0.766 mm
Heater element	Coiled Nichrome 22 AWG wire
Heater element pitch	7.62 mm
Fill valve	Nupro B-4HW bellows fill valve
Calorimeter	Coiled 3/16 in copper tubing

Figure 4. To calibrate the thermocouples, the chiller was set to various temperatures within the expected range of operation and allowed to reach steady state. At this point, the thermocouple readings were recorded. Polynomial curve fits were then used, which resulted in an uncertainty of $\pm 0.2^{\circ}\text{C}$ for all thermocouple readings.

The evaporator section consisted of a ceramic insulating material, Nichrome heating wire, a variable ac transformer, and power analyzers. The ceramic insulating paste was brushed on the evaporator section in several layers to prevent the Nichrome wire from coming in contact with the copper heat pipe. After the base layer of insulating material had dried, it was possible to wind the Nichrome heating wire (7.62 mm pitch) along the evaporator. Tape was used to hold the wire tight while another layer of ceramic was applied on top of the wire. The power analyzers consisted of an ammeter and voltmeter which measured delivered current and voltage, respectively. The uncertainty in the power reading was calculated to be $\pm 4\%$ of the reading.

To test the heat pipe, it was necessary to mount the heat pipe to a test stand. A Cenco Central Scientific Company optical test stand was used due to its availability and stability. Insulated clamps were used to isolate the heat pipe from any significant sources of heat conduction. To ensure a horizontal elevation of the heat pipe, a Schaevitz Angle Star Electronic Clinometer and adjustable legs were used. The stand was first leveled to the table then the heat pipe was leveled to the stand. After the heat pipe was instrumented, a 50-mm-thick layer of ceramic wool insulation and aluminum foil were tightly wrapped around the entire heat pipe to reduce heat losses to the environment.



- | | |
|----------------------|----------------------------|
| 1. Heat pipe | 6. Variable ac transformer |
| 2. Optical stand | 7. Calorimeter |
| 3. Electrical heater | 8. Flowmeter |
| 4. Ammeter | 9. Recirculating chiller |
| 5. Voltmeter | 10. Data acquisition |

Figure 3: Schematic of the heat pipe testing apparatus.

After straight testing was complete it was necessary to bend the pipe. Before the end caps were TIG welded to the container, the pipe was annealed to ensure a uniform bend. The prescribed annealing temperature for this alloy (C122000) was 375 to 650°C. Practice sections of pipe were used to ensure proper heat treatment and softening of the material. The pipe was then bent by hand around the edge of the centrifuge table without the outside surface kinking. Extra precautions had to be taken to ensure that the bent pipe was leveled with the centrifuge table. In this way, a direct comparison of the performance characteristics of the heat pipe in the straight and curved configurations was possible.

A typical test procedure is as follows. After the heat input to the evaporator section was set, the calorimeter temperature was adjusted until the center of the adiabatic section was stable within $\pm 0.1^\circ\text{C}$ of the target steady state adiabatic temperature (50 or 75°C). After the approximate target temperature in the adiabatic section was reached, the pipe was operated without interruption for at least 20 min. while the adiabatic temperature was monitored. If the adiabatic temperature did not change more than $\pm 0.1^\circ\text{C}$ from the set temperature, the pipe was considered to be at steady state and the temperature distribution for the pipe was recorded. Therefore, the adiabatic temperature was within $\pm 0.2^\circ\text{C}$ of the set point working temperature. The capillary limit was determined to be the heat input at which the evaporator end cap temperature increased significantly from the previous lower heat input. The end-to-end temperature drop divided by the heat input was defined as the thermal resistance ($R_{th} = (T_{eec} - T_{cec})/Q_e$).

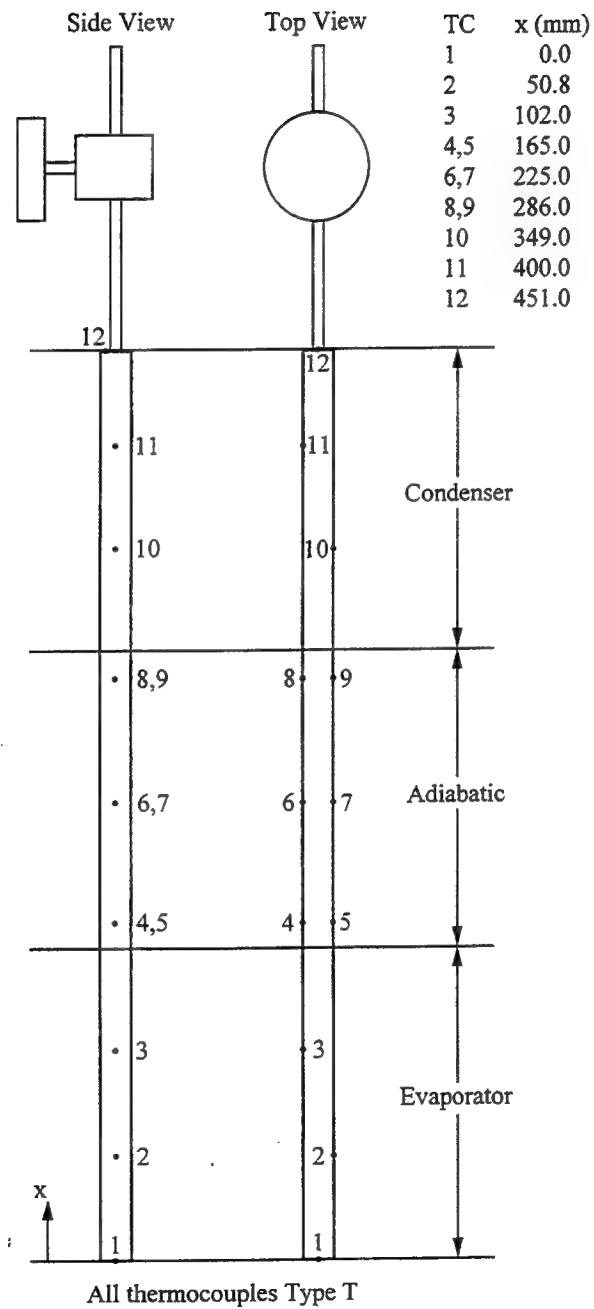


Figure 4: Thermocouple locations on the spirally-grooved heat pipe.

Results and Discussion

The goal of Phase 1 of this study was to experimentally determine the capillary limit and thermal resistance of a copper-ethanol spirally-grooved heat pipe in the straight and curved configurations. Figures 5 and 6 show the temperature distributions at the steady state for various heat inputs in the straight and curved configurations. At low power inputs, the temperature in the evaporator and adiabatic sections was nearly constant, while that in the condenser was slightly lower. As the heat input increased, the temperature in the evaporator began to increase significantly indicating a partial dryout situation. The evaporator temperature and length of the dry patch then increased with heat input beyond the point at which dryout initially began. The condenser temperature decreased with increasing heat input because the chiller temperature was decreased. Therefore, the thermocouples in the condenser section measured some average between the calorimeter and the heat pipe wall. Unfortunately, the two surface thermocouples were damaged during the disassembly process before the heat pipe was bent, so temperature information in the condenser section on the bent pipe was unavailable.

It was found that the capillary limit occurs at lower heat inputs for the curved heat pipe than in the straight configuration, as shown in Table 2. It is probable that the axial grooves were kinked and partially blocked, thereby significantly increasing the pressure drop in the liquid. Since the curvature of the pipe was extremely shallow, it is doubtful that the pressure drop in the vapor was affected. A planned post-mortem examination which will be undertaken after dynamic testing (Phase 2) so that the degree of groove kinking can be directly determined.

Table 2: Capillary limit results for the straight and curved spirally-grooved copper-ethanol heat pipe.

	Straight		Curved	
	50°C	75°C	50°C	75°C
Q_e (No Dryout)	25 W	30 W	15 W	15 W
Q_e (Partial Dryout)	30 W	35 W	20 W	20 W

Figure 7 shows the thermal resistance of the heat pipe in the straight and curved configurations. In general, the thermal resistance was fairly constant at the low heat inputs, but then increased as partial dryout in the evaporator occurred. The thermal resistance decreased with an increase in the working temperature of the pipe, due to the changes in the thermal properties at elevated temperatures. Figure 7 also shows that the resistance increased when the heat pipe was bent due to the kinking of the grooves.

The capillary limit results for the straight heat pipe were far from the analytical predictions, as shown in Figure 8. The analytical prediction assumes perfect wetting of the groove wall which is an ideal case with a considerably higher capillary limit. The most probable cause of the large decrease in the capillary limit is a possible contamination on the pipe wall surface which did not allow the wick to be properly wetted.

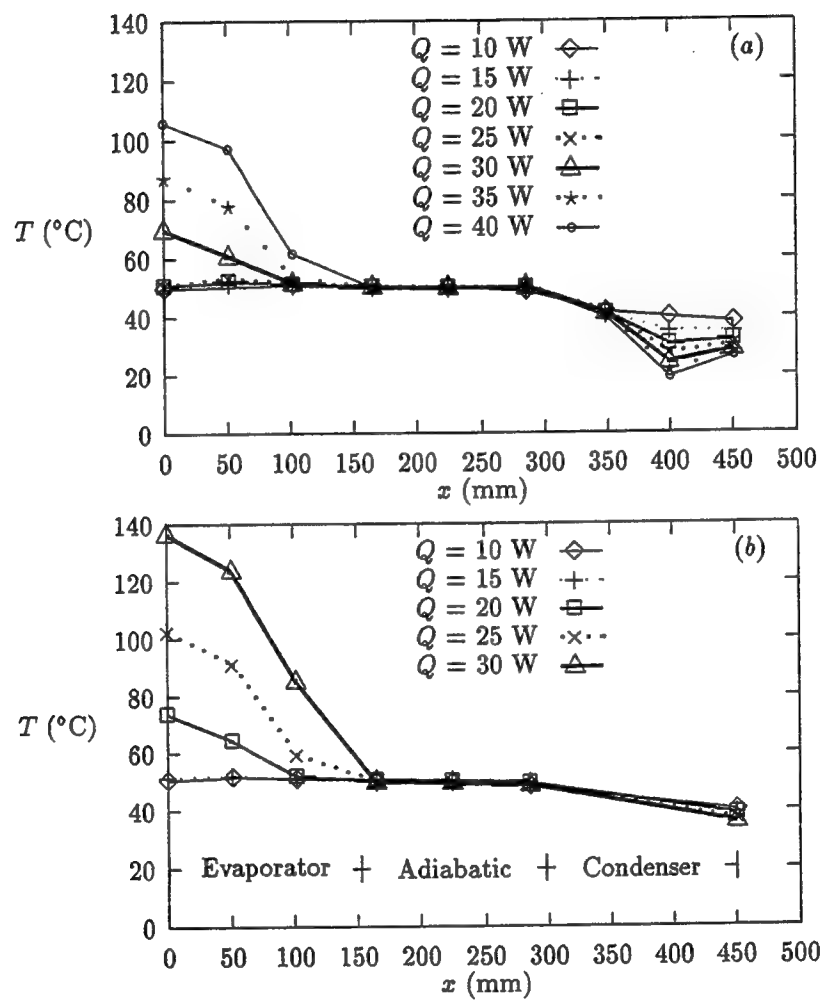


Figure 5: Temperature distributions for the spirally-grooved copper-ethanol heat pipe with $T_{\text{adiab}} = 50^\circ\text{C}$: (a) straight; and (b) curved.

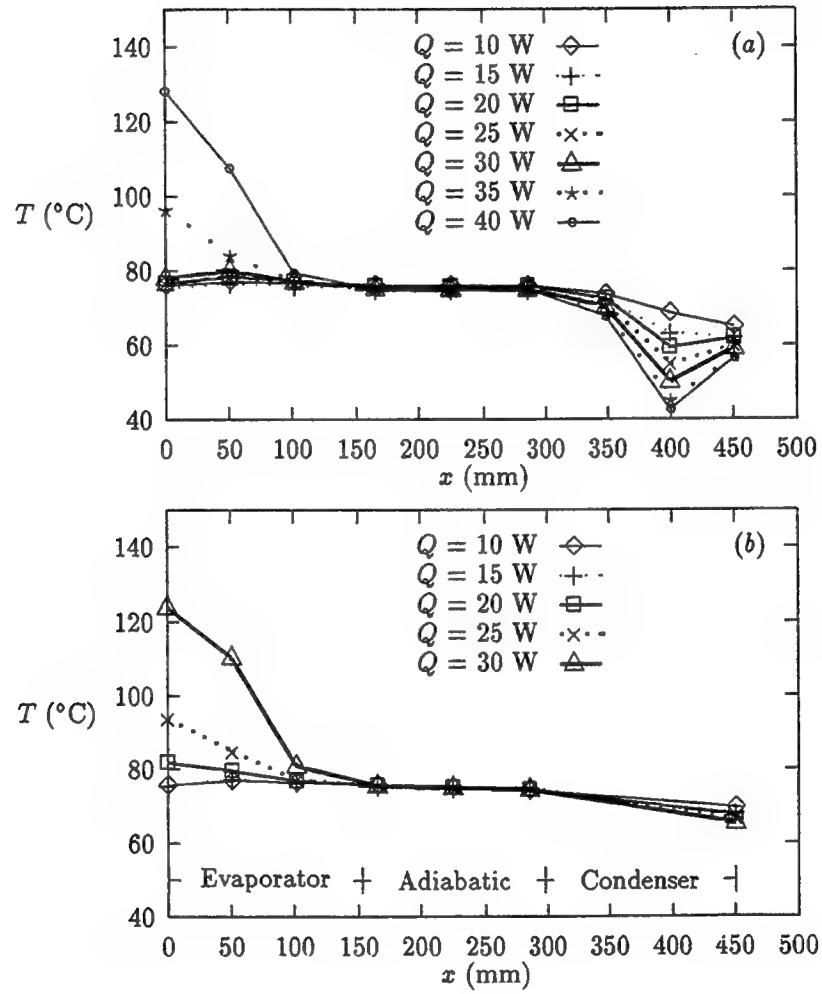


Figure 6: Temperature distributions for the spirally-grooved copper-ethanol heat pipe with $T_{\text{adiab}} = 75^\circ\text{C}$: (a) straight; and (b) curved.

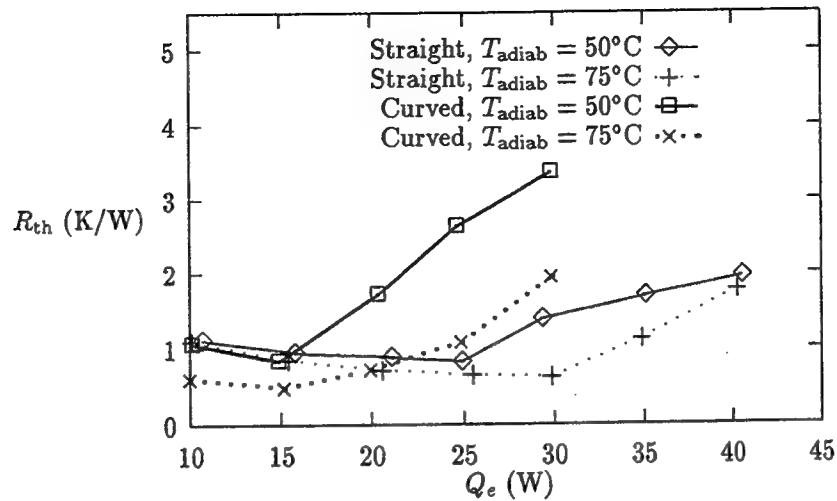


Figure 7: Thermal resistance of both the curved and straight configurations as functions of heat input.

Conclusions

The objective of the present experimental study was to determine the effect of pipe curvature on the performance of a spirally-grooved copper-ethanol heat pipe. It was found that the thermal resistance tends to be higher and the capillary limit lower for the curved configuration due to the fact that the grooves may have been kinked during the bending process. In addition, it was found that the capillary limit was overpredicted significantly by the analysis used, possibly due to contamination of the wick structure which prevented proper wetting characteristics. The second part of this study (Phase 2) will examine the effects of transient acceleration loading on the performance of the spirally-grooved heat pipe for possible application to thermal management of high-performance aircraft.

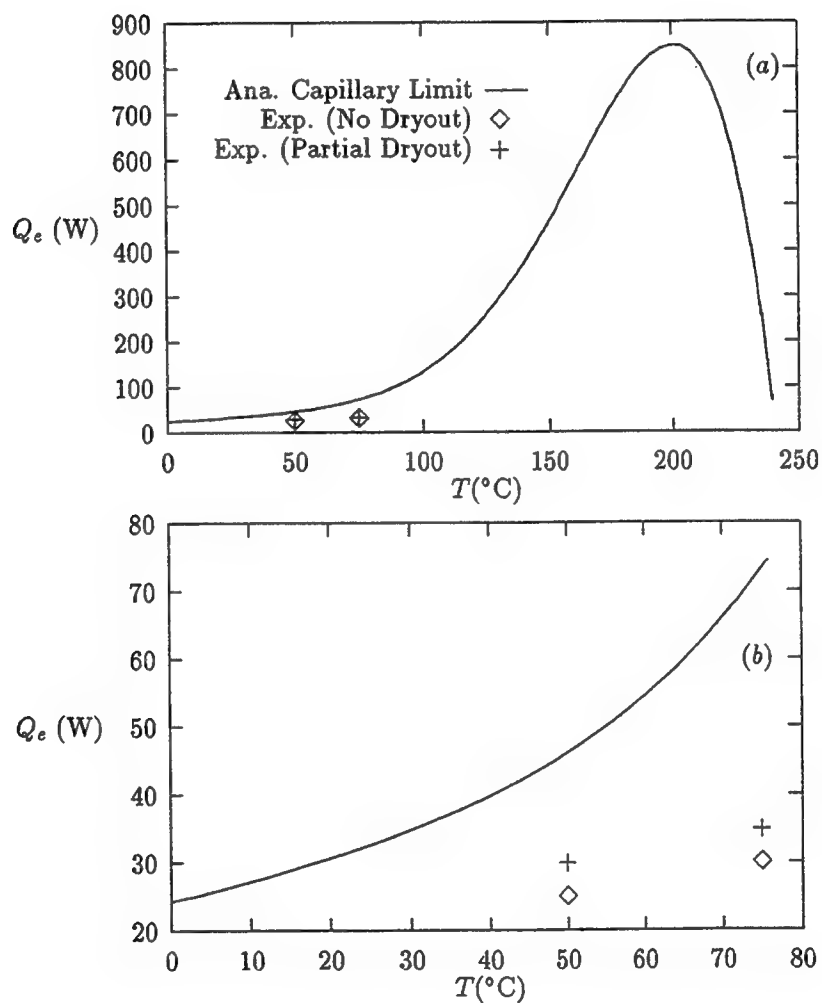


Figure 8: Capillary limit prediction and experimental data from the copper-ethanol straight heat pipe as a function of temperature: (a) $0 \leq T_{\text{adiab}} \leq 250^\circ\text{C}$; (b) $0 \leq T_{\text{adiab}} \leq 80^\circ\text{C}$.

Nomenclature

A	area, m^2
d_v	vapor core diameter, m
f	drag coefficient
F	friction coefficient, $(N/m^2)/W\text{-m}$
g	acceleration due to gravity, m/s^2
h_{fg}	latent heat of vaporization, J/kg
k	thermal conductivity, $W/(m\text{-K})$
K	permeability, m^2
L	length, m
n	number of grooves
P	pressure, N/m^2
Q	heat transfer rate, W
Q_{cm}	heat transfer rate at the capillary limit, W
Q_e	evaporator heat input, W
$(QL)_{cm}$	heat transport factor at the capillary limit, W-m
r	radius, m
R	principal radii of curvature, m
R_{th}	thermal resistance, K/W
Re	Reynolds number
T	temperature, K
U	radial vapor velocity, m/s
V	volume, m^3
w	groove width, m
\bar{w}	mean velocity, m/s
x	axial position, m
α	aspect ratio
δ	groove depth, m
ΔP	change in pressure, N/m^2
ϵ	porosity
μ	absolute viscosity, $kg/(m\text{-s})$
ρ_l	density, kg/m^3
σ	liquid surface tension, N/m
ψ	heat pipe inclination, rad
ω	angular velocity, rad/s

Subscripts

α	axial
adiab	adiabatic
c	capillary
cec	condenser end cap
circ	circumferential
cm	maximum capillary
cond	condenser
eec	evaporator end cap
evap	evaporator
h	hydraulic
in	inlet
l	liquid
min	minimum
p	pore
r	radial
ref	reference
t	total
th	thermal
v	vapor
w	wick
\perp	perpendicular

References

- Chi, S., 1976, *Heat Pipe Theory and Practice: A Sourcebook*, Hemisphere Publ. Corp., New York.
- Dunn, P., and Reay, D., 1982, *Heat Pipes*, 3rd Edn., Pergamon, Oxford.
- Gernert, N., et al., 1991, "Flexible Heat Pipe Cold Plates for Aircraft Thermal Control," *Proceedings of the Aerospace Technology Conference and Exposition*, SAE Paper No. 912105.
- Faghri, A., 1994, *Heat Pipe Science and Technology*, Taylor and Francis, Washington, DC.
- Faghri, A., and Parvani, S., 1988, "Numerical Analysis of Laminar Flow in a Double-Walled Annular Heat Pipe," *AIAA J. Thermophysics and Heat Transfer*, Vol. 2, No. 3, pp. 165-171.
- Shah, R.K., and Bhatti, M.S., 1987, "Laminar Convective Heat Transfer in Ducts," in *Handbook of Single-Phase Convective Heat Transfer*, Eds., Kakac, S., Shah, R.K., and Aung, W., Wiley, New York.

Thomas, S., and Yerkes, K., 1996, "Quasi-Steady State Thermal Resistance of a Flexible Copper-Water Heat Pipe Subjected to Transient Acceleration Loadings," *Proceedings of the ASME National Heat Transfer Conference*, AIChE Symposium Series Vol. 92, pp. 349-356.

Yerkes, K., and Beam, J., 1992, "Arterial Heat Pipe Performance in a Transient Heat Flux and Body Force Environment," *Proceedings of the Aerospace Atlantic Conference*, Dayton, OH, SAE Paper No. 921024.

Yerkes, K., and Hager, B., 1992, "Transient Response of Heat Pipes for Actuator Thermal Management," *Proceedings of the Aerospace Atlantic Conference*, Dayton, OH, SAE Paper No. 921024.

**SUBCRITICAL CRACK GROWTH OF Ti-6Al-4V
UNDER RIPPLE LOADING CONDITIONS**

**James P. Thomas
Assistant Professor
Department of Aerospace and Mechanical Engineering**

**University of Notre Dame
374 Fitzpatrick Hall
Notre Dame, IN 46556-5637**

**Final Report for:
Summer Faculty Research Program
Wright Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Wright Laboratory

August 1996

SUBCRITICAL CRACK GROWTH OF Ti-6Al-4V UNDER RIPPLE LOADING CONDITIONS

James P. Thomas
Assistant Professor
Department of Aerospace and Mechanical Engineering
University of Notre Dame

Abstract

Recent premature failures of some titanium alloy fan blades and disks in various military and commercial jet engines have been attributed to high cycle fatigue (HCF). The rotating components in jet engines experience "ripple loading"; that is, high frequency, low amplitude vibratory loads superposed on large, slowly varying "centrifugal" loads. An investigation of the influence of ripple-loading on subcritical crack growth of a mill-annealed Ti-6Al-4V titanium alloy was the object of this summer research. Experiments to characterize the: 1.) sustained load cracking threshold, $(K_{th})_{slc}$; 2.) the SLC growth rate as a function of K ; and 3.) the fatigue crack growth threshold at high R -ratios, ΔK_{th} were started. Testing preliminaries were completed. A d.c. potential drop-crack length calibration relationship was obtained using the data from two specimens along with initial estimates for $(K_{th})_{slc}$. The testing is being continued by WL/MLLN scientists.

SUBCRITICAL CRACK GROWTH OF Ti-6Al-4V UNDER RIPPLE LOADING CONDITIONS

James P. Thomas

Introduction

Recent failures of some titanium alloy fan blades and disks in various military and commercial jet engines have been attributed to high cycle fatigue (HCF). The failures occurred much earlier than predicted by the fatigue design calculations, and as a consequence, Air Force scientists are reassessing the problem of HCF damage and how it can be minimized or prevented.

The rotating components in jet engines are designed for operation in the "low cycle fatigue" (LCF) regime to account for the cyclic "centrifugal" loads experienced during power-up, cruise, and power-down operations. The rotating engine components are also exposed to higher frequency, lower amplitude vibratory loads due to aerodynamic interactions [1]. Resonant vibrations can occur in the fan or turbine stages during flight operation and are difficult to quantify. They result in the application of many cycles of stress of unknown magnitude to the rotating components. Fatigue damage accumulates very rapidly under these conditions resulting in crack initiation, crack growth, and eventual catastrophic fracture.

The superposition of a low amplitude, high frequency loads on top of a large mean load (stress ratio; $R \geq 0.9$) is known as "ripple loading" [2,3,4,5], a loading that is commonly experienced by rotating machinery. Ripple loading effects (RLE) in subcritical crack growth have been characterized as high R -ratio fatigue crack growth [2-5]. Yoder et al. [2] define a maximum stress intensity "window" for ripple-load cracking susceptibility. The upper bound corresponds to the threshold K for sustained load cracking or stress corrosion cracking (i.e., $K_{max} \leq (K_{th})_{slc}$ or $(K_{th})_{scc}$) and the lower bound corresponds to the threshold ΔK for fatigue crack growth (i.e., $K_{max} \geq \Delta K_{th}/(1 - R)$). In other words, a material is said to be susceptible to RLE when the applied maximum stress intensity is such that: $\Delta K_{th}/(1 - R) \leq K_{max} \leq (K_{th})_{slc}$ or $(K_{th})_{scc}$.

There is a need for better understanding of the influence of "ripple loading" on crack initiation and subcritical crack growth in titanium-based alloys. The work performed by Dr. Thomas of the University of Notre Dame this summer in the MLLN branch at Wright Laboratory focused on the experimental characterization of ripple loading effects on fatigue crack growth and sustained-load cracking of a Ti-6Al-4V alloy in a mill-annealed condition. A "short term" research plan was formulated and begun but not completed due to the time required for the crack growth rate tests. The work is being continued by WL/MLLN scientists.

The first phase of the research involves the characterization of: 1.) the sustained load cracking threshold, $(K_{th})_{slc}$; 2.) the SLC growth rate as a function of K ; and 3.) the fatigue crack growth threshold at high R -ratios, ΔK_{th} . This information will provide the K window for RLE. An important aspect of the above tasks will involve specification of threshold testing procedures. This is important because the loading history prior to the threshold measurement, as manifested by the monotonic and cyclic plastic zone sizes, can have a significant affect on the experimentally measured threshold K values [6,7]. It is also known that sustained load cracking and fatigue crack growth rate enhancement of the α and $\alpha+\beta$ titanium alloys in air are due to the influence of internal hydrogen [8,9,10,11,12,13]. It is therefore likely that the hydrogen and

hydride distributions in the crack tip region just prior to threshold measurement will have an influence on the measured value.

The crack growth rate testing is being conducted on mini-C(T) specimens machined from Ti-6Al-4V forged bar stock used for fan blades. A d.c. potential drop technique is used for making indirect crack length measurements. The tests are conducted on an MTS servohydraulic load frame with computer automated control and data acquisition. Most of the tests are being conducted in laboratory air at room temperature; a few exploratory tests at room temperature in a vacuum atmosphere are planned to determine the influence of external sources of hydrogen on the subcritical cracking process.

Three series of tests must be completed before the ripple loading tests can be started. These include tests for: 1.) obtaining data for the d.c. potential crack length-voltage calibration relationship; 2.) determining $(K_{th})_{slc}$, the threshold K for sustained-load cracking; and 3.) determining ΔK_{th} , the threshold ΔK for fatigue crack growth. After these tests have been completed, tests will be conducted to see if ripple-loading fatigue can be treated simply as a case of high R -ratio fatigue, or whether there are additional interactions with the sustained load cracking process. One definite planned experiment is to attempt to duplicate a result found by Marci [14] for an IMI 834 alloy. He found that da/dN actually increased with decreasing ΔK , the increase starting close to ΔK_{th} , when $K_{max} \geq 30 \text{ MPa}\sqrt{m}$. Questions to be answered include: does this behavior occur with the Ti-6-4 alloy; can it be predicted using a superposition model which relates the total crack growth rate to the SLC and fatigue crack growth rates; and are there loading history effects that play a role in this phenomenon?

The calibration tests have been completed and crack length/voltage data from two specimens has been used to determine an empirical calibration relationship. A preliminary estimate for $(K_{th})_{slc}$ was obtained at the end of the calibration tests; the results indicate a $(K_{th})_{slc}$ value somewhere in the range of 21.5 to 25 $[\text{MPa}\sqrt{m}]$.

The first specimen in the second series of tests to determine $(K_{th})_{slc}$ was started before Dr. Thomas left for the summer. The test procedure consisted of increasing K from 20.0 in 0.5 $\text{MPa}\sqrt{m}$ increments while watching for crack growth. A minimum of 6 $[hr]$ hold time between K increases was used to check for crack growth. The crack had still not shown any evidence of growth with $K = 30 \text{ MPa}\sqrt{m}$. A large amount of plastic deformation at the crack tip indicated blunting probably due to room temperature creep, possibly with some influence from internal hydrogen [15,16]. Internal hydrogen levels in the calibration specimens and in this specimen are being determined to rule out the possibility of different hydrogen levels as being a factor in the threshold differences. Work continues on developing a test procedure to give valid $(K_{th})_{slc}$ values while properly accounting for the prior load history.

Material and Experimental Work

Mini-compact tension C(T) specimens with $W = 20$ and $B = 4.9 \text{ [mm]}$ and C-R crack plane orientation (see Figure 1) were machined from a 44.5 $[\text{mm}]$ diameter forged bar of Ti-6Al-4V. The post-forging heat treatment of the bar consisted of vacuum annealing for 2 hours at 1300 $^{\circ}\text{F}$, Argon gas quench, followed by a stress relief treatment at 1020 $^{\circ}\text{F}$ for 2 hours, and a final Argon gas quench. The specimen faces were polished to a mirror finish to aid the optical crack

length measurements. The chemical composition and some properties of the material are given in Tables 1 and 2.

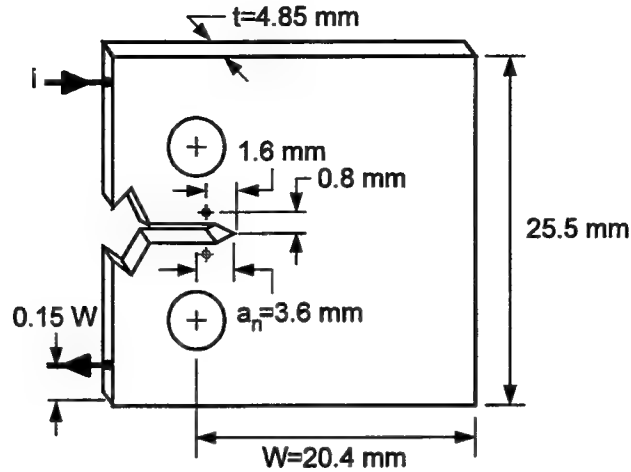


Figure 1: Mini-compact tension specimen geometry and DCEP lead locations.

Chemical Composition (Wgt. %)							
Al	V	O	Fe	C	N	H	Ti
6.19	4.11	0.174	0.146	0.030	0.005	0.0027	Bal.

Table 1: Chemical composition of the Ti-6Al-4V test material.

Material Properties			
S_{ult} [MPa]	S_y [MPa] (0.2%)	% RA	Conductivity [$1/\Omega\cdot m$]
998	927	44.1	585,000

Table 2: Material properties for the Ti-6Al-4V test material.

Crack length was measured using a d.c. electric potential (DCEP) technique; optical measurements were made on both sides of the specimen to confirm the DCEP measurements. The DCEP leads were spot welded on opposite sides of the specimen straddling the notch. The potential and current lead locations are shown in Figure 1.

Two specimens were used to obtain data for an empirical crack-length/potential drop calibration relationship. The tests were performed at room temperature in laboratory air using a 22 [kN] MTS servohydraulic load frame with the Mate¹ computer control and data acquisition software. Use was made of an approximate crack length-potential drop relationship² in order to run “constant K_{max} ” calibration tests. The expression used for the stress intensity factor is that recommended for C(T) specimens in ASTM E-647-93:

$$K = \frac{P}{B\sqrt{W}} \frac{(2 + a/W)}{(1 - a/W)^{3/2}} \times \left(0.886 + 4.64\left(\frac{a}{W}\right) - 13.32\left(\frac{a}{W}\right)^2 + 14.72\left(\frac{a}{W}\right)^3 - 5.6\left(\frac{a}{W}\right)^4 \right) \quad (1)$$

¹ Developed and maintained by the University of Dayton Research Institute, Dayton, OH.

² The accuracy of the approximate relationship is estimated to be better than $\pm 2\%$.

The specimens were fatigue precracked a minimum of 1 [mm] beyond the tip of the notch at $K_{max} = 10 [MPa\sqrt{m}]$ and $R = 0.1$ or 0.7 . The crack front was demarcated during calibration by keeping K_{max} = constant and alternating between high and low R values. Figure 2 shows the well marked fracture surfaces of the two calibration specimens.

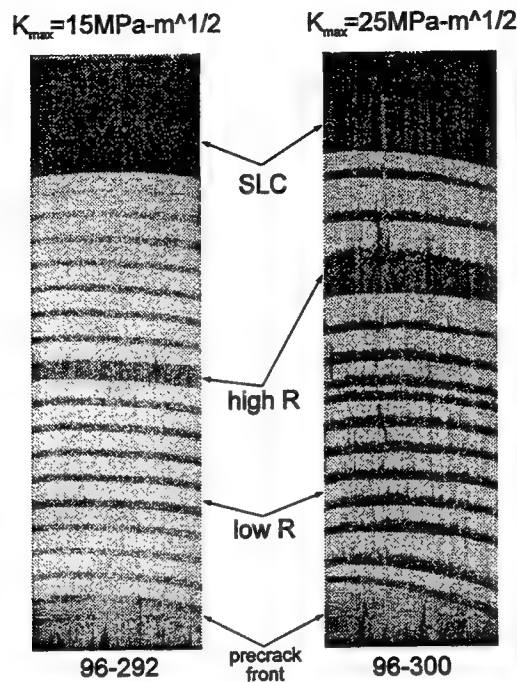


Figure 2: The fracture surfaces of the two calibration specimens showing the distinct marking that comes about by varying the R -ratio.

The first calibration test (test specimen # 96-300) was run at $K_{max} = 25 [MPa\sqrt{m}]$ and $R = 0.9$ & 0.5 at the start of the test, and $R = 0.85$ & 0.55 in later in the test (to speed up the rate of crack growth). A total of 30 crack-length/potential drop data points were obtained. The second test (test specimen # 96-292) was run at $K_{max} = 15 [MPa\sqrt{m}]$ and $R = 0.8$ & 0.2 at the start of the test, and $R = 0.75$ & 0.2 later in the test (to speed up the rate of crack growth). A total of 38 crack-length/potential drop data points were obtained for this test.

The potential drop system was consisted of an HP-6033A power supply operated in a constant current mode and an HP-3478A multimeter; both were controlled by the Mate software using an GPIB connection. A current of 5 [A] was used during the tests, and the full-scale range on the voltmeter was set at 300 [mV]. Potential measurements were made with and without the current flowing. The potential at 0 [A] corresponds to the thermal voltage which is subtracted from the 5 [A] potential measurement to give an actual specimen voltage drop across the potential probes.

Crack lengths were measured at the end of each test using a Gaertner toolmaker's microscope with digital readout. Duplicate measurements, on each half of the fractured specimen, were made at five equally spaced locations along the crack front. Each crack length value used in the analysis of the calibration relationship therefore consisted of an average of 20 independent crack length measurements.

Some rough estimates of the sustained load cracking threshold values, $(K_{th})_{slc}$, were obtained at the end of each calibration test. Both specimens had rather large a/W values (> 0.84) at this point in testing which made precise K control difficult. In the first test (96-300), the stress ratio was set to $R = 1.0$ and K_{max} was kept at $25 [MPa\sqrt{m}]$; in the second test (96-292), the stress ratio was set to $R = 1.0$ and K_{max} was increased from 15 to $25 [MPa\sqrt{m}]$. The crack grew under this sustained load level in both specimens. The first specimen failed before the test could be stopped, but in the second test, K_{max} was decreased from $25 [MPa\sqrt{m}]$ in steps of $0.5 [MPa\sqrt{m}]$ until the crack arrested at $K_{max} = 21.5 [MPa\sqrt{m}]$.

Results and Discussion

A normalized calibration relationship is used in the testing. This development of this relationship is shown below. From Ohm's law:

$$V = IR \quad (2)$$

where

$$R = \rho \frac{l}{A} \quad (3)$$

with $\rho :=$ resistivity ($= 1/\kappa$; $\kappa :=$ conductivity [$1/\Omega\text{-}m$]); $l :=$ "length" between the potential leads; $A :=$ "cross-sectional area" between the potential leads. Now take:

$$\frac{l}{A} = \frac{f(\frac{a}{W})}{B} \quad (4)$$

where $f(\frac{a}{W}) :=$ non-dimensional function of the normalized crack length, a/W , and $B :=$ specimen thickness. From the above relations, we obtain the relationship:

$$f(\frac{a}{W}) = \frac{\kappa BV}{I} \quad (5)$$

which is inverted to obtain normalized crack length as a function of normalized potential:

$$\frac{a}{W} = f^{-1}\left(\frac{\kappa BV}{I}\right) := g\left(\frac{\kappa BV}{I}\right) \quad (6)$$

Variations in material conductivity, specimen thickness, and supply current are accounted for by the voltage normalization. The exact form of $g(\frac{\kappa BV}{I})$ is unique to each particular specimen geometry, and potential and current lead placement location. In this work, we fit a cubic polynomial to the set of calibration data $\{(\frac{a}{W})_i, (\frac{\kappa BV}{I})_i\}$ ($i = 1 \dots n = 67$). All data used in the analysis has been included in the Appendix. A least squares "error-in-variables" procedure³ is used to perform the regression analysis. The error-in-variables technique differs from classical analysis in that the independent variables (i.e., the $\frac{\kappa BV}{I}$ values) are not assumed to be known without measurement error.

³ See Thomas & Wei [17]; the Fortran programs used to perform the analysis are included in the Appendix.

Figure 3 shows a scatter plot of calibration data with the error-in-variables curve for $g\left(\frac{\kappa BV}{I}\right)$. Figure 4 shows the corresponding residuals for the curve fit. The empirical calibration relationship determined by the above analysis is given by:

$$\begin{aligned} \frac{a}{W} = & 0.13106 \times 10^{-1} + 0.23778 \left(\frac{\kappa BV}{I} \right) \\ & + 0.21020 \times 10^{-1} \left(\frac{\kappa BV}{I} \right)^2 - 0.66706 \times 10^{-2} \left(\frac{\kappa BV}{I} \right)^3 \end{aligned}$$

Variation in the placement of the potential leads on each specimen leads to errors in the crack length measurement. These errors can be decreased by shifting the measured potential drop by some constant offset value. The calibration relationship is redefined as:

$$\frac{a}{W} = g\left(\frac{\kappa B}{I}(V - V_{offset})\right) \quad (7)$$

where V_{offset} is determined using the measured potential $V\left(\frac{a}{W}\right)$ at some known a/W value, and the inverse calibration relationship expressed by Eq. (5):

$$V_{offset} := V\left(\frac{a}{W}\right) - \left(\frac{I}{\kappa B}\right)f\left(\frac{a}{W}\right) \quad (8)$$

This shifting procedure causes $(a/W)_{pred} = (a/W)_{meas}$ at the point where V_{offset} is calculated.

The function $f\left(\frac{a}{W}\right)$ is also determined by least squares, error-in-variables fitting of a polynomial to the data $\left\{\left(\frac{\kappa BV}{I}\right)_i, \left(\frac{a}{W}\right)_i\right\}$ ($i = 1 \dots n = 67$). The resulting expression is given by:

$$\frac{\kappa BV}{I} = \left(\frac{a/W}{1 - a/W}\right) \left\{ 3.8175 - 3.9743 \left(\frac{a}{W}\right) + 0.33187 \left(\frac{a}{W}\right)^2 \right\} \quad (9)$$

Figure 5 shows a scatter plot of the data with the curve given by Eq. (9), and Figure 6 shows the corresponding residuals.

Figures 7 and 8 show scatter and residual plots for the calibration data shifted using each specimen's notch data (length & potential) to determine V_{offset} . A slight reduction in error is evident. The potential drop at the notch is highly sensitive to machining and lead placement errors, and is probably not the best point to use in making the calibration correction.

Figure 9 shows fractographs taken from various regions on the first calibration specimen (#96-300; $K_{max} = 25 \text{ MPa}\sqrt{m}$). There are noticeable differences between the low and high R regions, and between the fatigue and sustained load regions. These differences, unfortunately, cannot be seen in the printed figures. Striations can be seen on the low R surface but not on the high R surface. The sustained load region shows larger cleavage facets than the fatigue; the cleavage facets on the SLC and high R fatigue surfaces exhibit many small tear ridges.

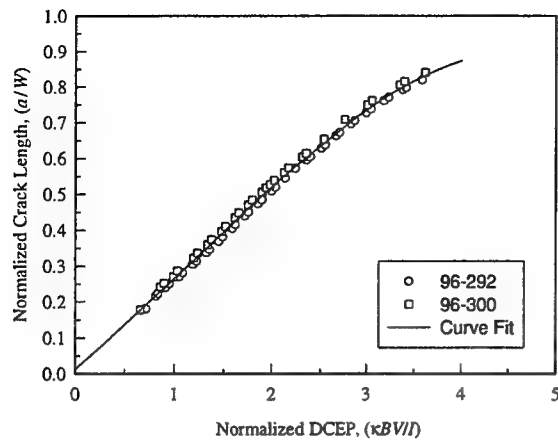


Figure 3: Scatter plot for the calibration relationship $g\left(\frac{\kappa BV}{I}\right)$.

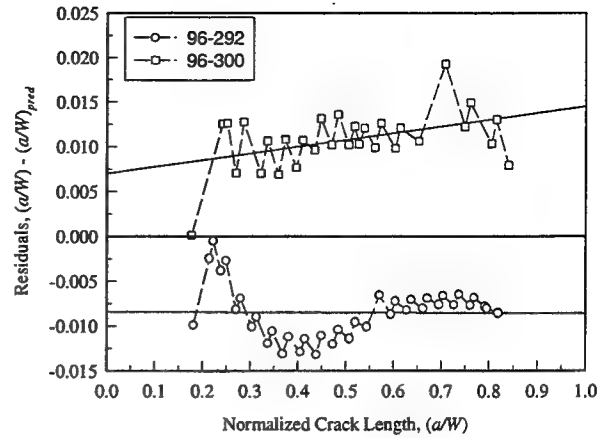


Figure 4: Residuals for the curve fit in Fig. 3.

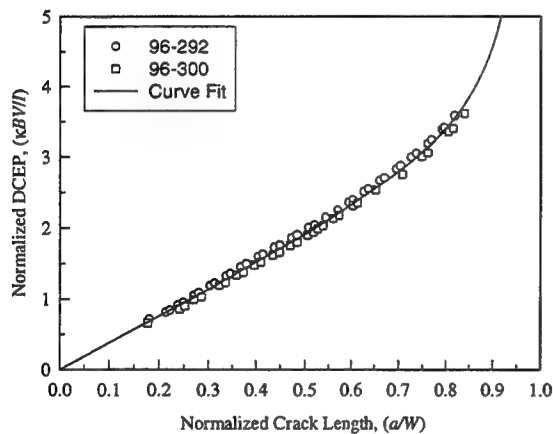


Figure 5: Scatter plot for the calibration relationship $f\left(\frac{a}{W}\right)$.

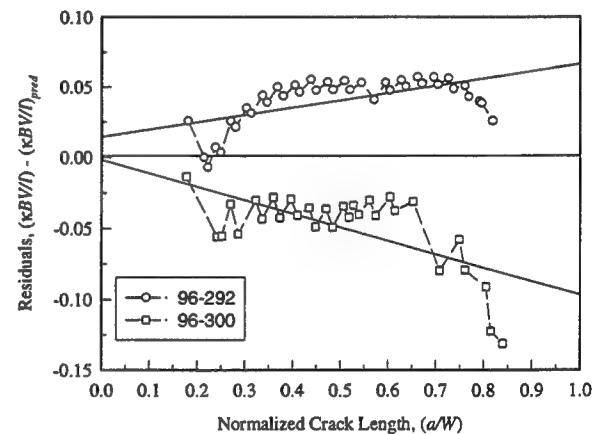


Figure 6: Residuals for the curve fit in Fig. 5.

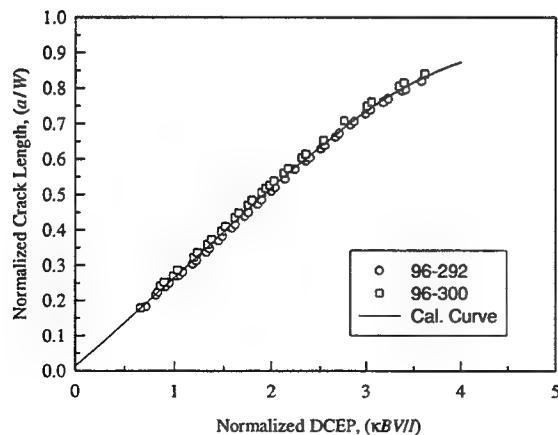


Figure 7: Scatter plot for the "notch-potential" shifted calibration data.

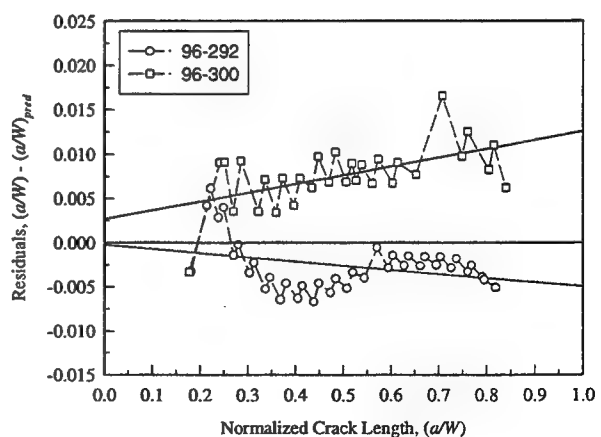
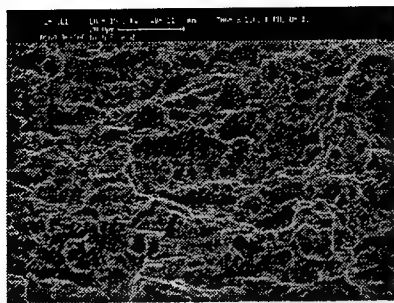
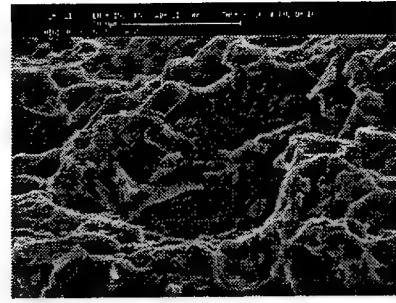


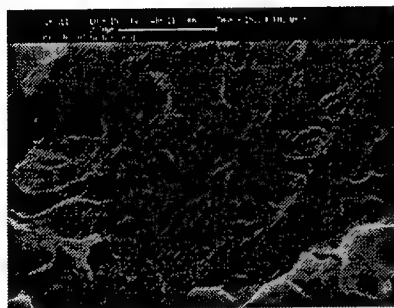
Figure 8: Residuals for the curve fit in Fig. 7.



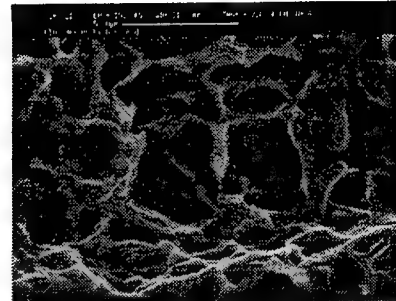
(a)



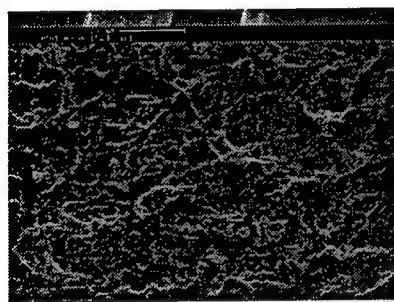
(b)



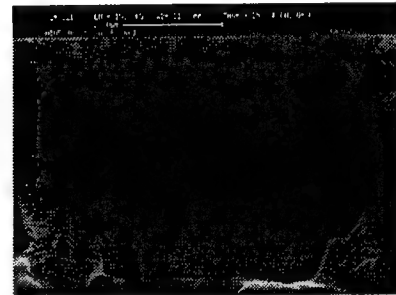
(c)



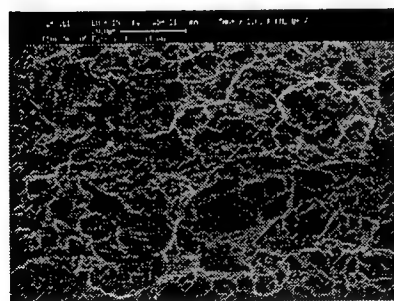
(d)



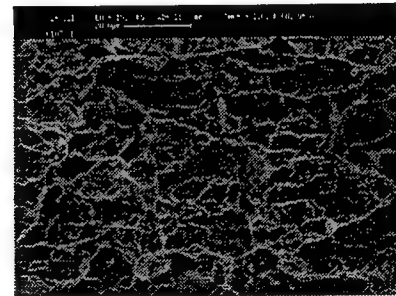
(e)



(f)



(g)



(h)

Figure 9: Fractographs of the calibration specimen #96-300. Figures (a)-(d) correspond to high R fatigue; (e)-(f) to low R fatigue; (g) to the transition zone between fatigue and sustained load cracking (slc); and (h) to slc. (Fractography by Dr. Y. Lenets).

Summary

An investigation to determine the influence of ripple loading on subcritical crack growth of a mill-annealed Ti-6Al-4V titanium alloy was started. Experiments were planned to characterize the: 1.) sustained load cracking threshold, $(K_{th})_{slc}$; 2.) the SLC growth rate as a function of K ; and 3.) the fatigue crack growth threshold at high R -ratios, ΔK_{th} . Testing preliminaries were completed. A d.c. potential drop-crack length calibration relationship was obtained as well as preliminary estimates for $(K_{th})_{slc}$ in the range of 21.5 to 25 $[MPa\sqrt{m}]$. The testing is being continued by WL/MLLN scientists.

Acknowledgments

I want to thank AFOSR and Wright Laboratory for the financial support for my summer visit at WL/MLLN through the Summer Faculty Research program. I also want to thank Dr. Ted Nicholas, my laboratory mentor, for inviting me to spend the summer with the MLLN group, for many interesting and enjoyable discussions, and for his research and career advice. Finally, I must thank Drs. Youri Lenets, Andy Rosenberger, and George Sendeckyj for their help and companionship during this summer. They made my stay in Dayton both productive and pleasant.

References

1. Powell, B. E., T. V. Duggan, and R. Jeal, "The Influence of Minor Cycles on Low Cycle Fatigue Crack Propagation", *Int. J. Fatigue*, Jan. 1982, pp. 4-14.
2. Yoder, G. R., P. S. Pao, and R. A. Bayles, "Ripple-Load Cracking in a Titanium Alloy", *Scripta MET.*, Vol. 24, 1990, pp. 2285-2289.
3. Pao, P. S., R. A. Bayles, and G. R. Yoder, "Effect of Ripple Load on the Stress-Corrosion Cracking in Structural Steels", Paper # OTC 5983 in *Offshore Technology Conference*, 1989, pp. 245-252.
4. P. S. Pao, D. A. Meyn, R. A. Bayles, C. R. Feng, and G. R. Yoder, "Effect of Ripple Loads on Sustained-Load Cracking in Titanium Alloys", *Scripta MET.*, Vol. 33, 1995, pp. 497-501.
5. Horstmann, M. and J. K. Gregory, "Observations on the Ripple Loading Effect", *Scripta MET.*, Vol. 25, 1991, pp. 2503-2506.
6. Döker, H., and M. Peters, "Fatigue Threshold Dependence on Material, Environment, and Microstructure", in *Fatigue 84*, C. J. Beevers, Ed., 1984, pp. 275-285.
7. Marci, G., D. E. Castro, and V. Bachmann, "Closure Measurements via a Generalized Threshold Concept", in *Fracture Mechanics: 21st Symposium*, ASTM STP 1074, J. Gudas, J. A. Joyce, and E. M. Hackett, Eds., American Society for Testing and Materials, Philadelphia, PA, 1990, pp. 563-580.
8. Meyn, D., "Effects of Very Low Pressures on Fatigue Fracture Surface Features in Ti-6Al-4V", in *Fractography-Microscopic Cracking Processes*, ASTM STP 600, American Society for Testing and Materials, Philadelphia, PA, 1976, pp. 75-87.
9. Moody, N. R., and J. E. Costa, "A Review of Microstructural Effects on Hydrogen-Induced Sustained Load Cracking in Structural Titanium Alloys", in *Microstructure/Property Relationships in Titanium Aluminides and Alloys*, Y. W. Kim and R. R. Boyer, Eds., Minerals, Metals, & Materials Society, 1991, pp. 587-604.
10. Yoder, G. R., C. A. Griffis, and T. W. Crooker, "The Cracking of Ti-6Al-4V Alloys Under Sustained Load in Ambient Air", *Trans. ASME: J. Eng. Matls. & Tech.*, Oct. 1974, pp. 268-274.
11. Williams, D. N., "Effects of Hydrogen in Titanium Alloys on Subcritical Crack Growth under Sustained Load", *Matls. Sci. & Eng.*, Vol. 24, 1976, pp. 53-63.
12. Meyn, D. A., "Effect of Hydrogen on Fracture and Inert-Environment Sustained Load Cracking Resistance of α - β Titanium Alloys", *Met. Trans. A*, Vol. 5, 1974, pp. 2405-2414.
13. Meyn, D. A., "Temperature Dependence of Sustained Load Cracking Caused by Residual Hydrogen in Ti-6Al-4V", in Vol. 4 of *Titanium, Science, and Technology*, G. Lütjering, U. Zwicker, W. Bunk, Eds., Germany, 1985, pp. 2511-2518.
14. Marci, G., "Failure Mode Below 390K with IMI 834", in *Fatigue 1996*, 1996, pp. 493-498.
15. Neal, D. F., "Creep Fatigue Interactions in Titanium Alloys", in *Sixth World Conference on Titanium*, Part 1, P. Lacombe, R. Tricot, and G. Béranger, Eds., Société Française de Métallurgie, France, 1988, pp. 175-180.

16. Yin, H., M. Gao, and R. P. Wei, "Deformation and Subcritical Crack Growth under Static Loading", *Matls. Sci. and Eng.*, Vol. A119, 1989, pp. 51-58.
17. Thomas, J. P., and R. P. Wei, "Standard-Error Estimates for Rates of Change fro Indirect Measurements", *TECHNOMETRICS*, Vol. 38, 1996, pp. 59-68.

APPENDIX

DCEP Calibration Data for the Mini-C(T) Specimen

Material:	TI-6Al-4V
Specimen ID:	96-292
Thickness [m]:	0.0048006
Width [m]:	0.020412
Current [A]:	5
Conductivity [1/ohm-m]:	585000
Specimen ID:	96-300
Thickness [m]:	0.0048387
Width [m]:	0.020441
Current [A]:	5
Conductivity [1/ohm-m]:	585000

Notes:

1. Analysis consists of LS fitting a cubic polynomial to the (a/W) -vs- (kBV/I) data.
2. Obs. numbers 1 through 38 plus 39* were obtained from 96-292; Obs. numbers 39 through 68 were obtained from 96-300.
3. Notch data for 96-292 corresponds to Obs. No. 1; Notch data for 96-300 corresponds to Obs. No. 39.

Test Data

Obs. No./Spec. No.	Crack Length, a [m]	a/W [1]	DCEP, v [volts]	Max Load, [kN]	Load Ratio, R	Kmax [MPa-m ^{1/2}]	DelK [MPa-m ^{1/2}]
1/96-292	0.003695	0.181018	0.00126944	1.7028	0.798	10.01	2.02
2/96-292	0.004387	0.214914	0.00145190	1.583	0.108	10.30	9.19
3/96-292	0.004550	0.222916	0.00149380	1.551	0.701	10.33	3.09
4/96-292	0.004864	0.238296	0.00162200	1.473	0.101	10.24	9.21
5/96-292	0.005096	0.249670	0.00169240	2.145	0.802	15.39	3.05
6/96-292	0.005526	0.270743	0.00187400	2.022	0.204	15.35	12.22
7/96-292	0.005725	0.280468	0.00193230	1.957	0.799	15.24	3.06
8/96-292	0.006203	0.303893	0.00211460	1.844	0.205	15.27	12.14
9/96-292	0.006397	0.313387	0.00217220	1.788	0.801	15.17	3.02
10/96-292	0.006877	0.336900	0.00235420	1.681	0.206	15.15	12.03
11/96-292	0.007072	0.346444	0.00241070	1.641	0.798	15.15	3.06
12/96-292	0.007526	0.368718	0.00258200	1.549	0.207	15.15	12.01
13/96-292	0.007772	0.380776	0.00265290	1.505	0.747	15.19	3.84
14/96-292	0.008270	0.405141	0.00283440	1.407	0.208	15.14	11.99
15/96-292	0.008474	0.415139	0.00289470	1.373	0.749	15.17	3.81
16/96-292	0.008955	0.438695	0.00307330	1.290	0.203	15.20	12.11
17/96-292	0.009177	0.449571	0.00313550	1.248	0.751	15.16	3.77
18/96-292	0.009662	0.473369	0.00331320	1.165	0.210	15.15	11.97
19/96-292	0.009893	0.484674	0.00338340	1.131	0.752	15.20	3.77
20/96-292	0.010381	0.508554	0.00356510	1.055	0.213	15.26	12.01
21/96-292	0.010608	0.519678	0.00363370	1.019	0.751	15.27	3.80
22/96-292	0.011105	0.544024	0.00381960	0.940	0.210	15.26	12.06
23/96-292	0.011656	0.571040	0.00399860	0.835	0.784	14.90	3.22
24/96-292	0.012136	0.594571	0.00419880	0.779	0.219	15.18	11.86
25/96-292	0.012329	0.604015	0.00426200	0.738	0.785	14.93	3.21
26/96-292	0.012825	0.628324	0.00446620	0.689	0.216	15.40	12.07
27/96-292	0.013017	0.637731	0.00453430	0.642	0.783	14.94	3.24
28/96-292	0.013513	0.662034	0.00474700	0.576	0.225	14.97	11.60
29/96-292	0.013701	0.671205	0.00481650	0.554	0.775	15.04	3.38
30/96-292	0.014197	0.695520	0.00503820	0.496	0.217	15.22	11.92
31/96-292	0.014382	0.704566	0.00511170	0.469	0.786	15.10	3.53
32/96-292	0.014840	0.727008	0.00533300	0.415	0.206	15.17	12.04
33/96-292	0.015057	0.737648	0.00542560	0.388	0.748	15.11	3.81
34/96-292	0.015530	0.760824	0.00567450	0.347	0.189	15.67	12.80
35/96-292	0.015713	0.769784	0.00576130	0.310	0.748	14.88	3.75
36/96-292	0.016165	0.791952	0.00602270	0.274	0.179	15.46	12.69
37/96-292	0.016259	0.796562	0.00608020	0.254	0.683	14.85	4.71
38/96-292	0.016723	0.819260	0.00637570	0.220	0.211	15.52	12.24
39/96-300	0.003630	0.177584	0.00116703	1.7345	0.8005	10.01	2.00
40/96-300	0.004935	0.241444	0.00151920	1.5100	0.105	10.50	9.40
41/96-300	0.005142	0.251571	0.00158770	3.6440	0.801	26.05	5.18
42/96-300	0.005517	0.269912	0.00175020	3.3100	0.099	24.86	22.40
43/96-300	0.005848	0.286084	0.00182150	3.3240	0.9	26.05	2.60
44/96-300	0.006591	0.322424	0.00210750	3.0170	0.499	25.97	13.01
45/96-300	0.006869	0.336024	0.00217560	2.9120	0.9	25.96	2.60
46/96-300	0.007347	0.359404	0.00236070	2.7500	0.499	26.03	13.04
47/96-300	0.007630	0.373283	0.00242910	2.6500	0.9	26.00	2.60
48/96-300	0.008109	0.396694	0.00261150	2.5040	0.5	26.12	13.06
49/96-300	0.008384	0.410164	0.00268340	2.4130	0.85	26.08	3.91
50/96-300	0.008885	0.434674	0.00286100	2.2720	0.551	26.25	11.78
51/96-300	0.009160	0.448100	0.00293030	2.1980	0.85	26.36	3.95
52/96-300	0.009606	0.469952	0.00310450	2.0660	0.552	26.37	11.81
53/96-300	0.009885	0.483602	0.00317740	1.9880	0.853	26.41	3.88
54/96-300	0.010338	0.505732	0.00335980	1.8690	0.554	26.56	11.85
55/96-300	0.010581	0.517636	0.00343100	1.7900	0.851	26.41	3.94
56/96-300	0.010763	0.526521	0.00350970	1.7540	0.552	26.63	11.93
57/96-300	0.010999	0.538090	0.00358150	1.6830	0.849	26.55	4.01
58/96-300	0.011457	0.560494	0.00376340	1.5560	0.556	26.51	11.77
59/96-300	0.011720	0.573379	0.00383960	1.4970	0.85	26.72	4.01
60/96-300	0.012337	0.603537	0.00409130	1.3310	0.558	26.64	11.77
61/96-300	0.012547	0.613814	0.00415360	1.2850	0.848	26.80	4.07
62/96-300	0.013341	0.652670	0.00447600	1.0970	0.561	27.06	11.88
63/96-300	0.014469	0.707847	0.00486820	0.8720	0.849	28.34	4.28
64/96-300	0.015314	0.749164	0.00531290	0.6770	0.56	28.09	12.36
65/96-300	0.015554	0.760925	0.00540100	0.6130	0.861	27.46	3.82
66/96-300	0.016463	0.805373	0.00592050	0.4450	0.56	27.67	12.18
67/96-300	0.016665	0.815270	0.00600440	0.4050	0.855	27.36	3.97
68/96-300	0.017173	0.840135	0.00638510	0.3200	0.557	27.16	12.03

```

C      PROGRAM CAL1
C
C      MODIFIED FOR DCEP CALIBRATION ANALYSIS; 16-JUL-95.
C      3RD ORDER POLYNOMIAL FIT OF:  $(a/W) - us. - (\kappa Bv/I)$ 
C      WITH ALL WEIGHTING FACTORS SET EQUAL TO 1.0.
C
C      PROGRAM FOR ERROR-IN-VARIABLES, MULTI-LINEAR LEAST
C      SQUARES CURVE FITTING USING IMSL ROUTINES.
C
C      BOTH THE DEPENDENT AND INDEPENDENT VARIABLES HAVE SOME
C      MEASUREMENT ERROR ASSOCIATED WITH THEM. THE ANALYSIS
C      IS DESCRIBED BY W. H. SOUTHWELL, "THE COMPUTER JOURNAL",
C      VOL. 19, NO. 1, 1974, PP. 69-73.
C
C      INSTRUCTIONS FOR USE: THE PROGRAM CURRENTLY WRITTEN TO FIT
C      A CUBIC POLYNOMIAL  $Y=B0+B1*X+B2*X**2+B3*X**3$  TO THE DATA SET
C       $\{Y(I), X(I)\}$ ,  $I=1, \dots, NOBS$ , WHERE THE  $Y(I)$  AND THE  $X(I)$  ARE
C      SUBJECT TO UNCERTAINTY DUE TO MEASUREMENT ERRORS.
C      1. SET NOBS=NUMBER OF DATA POINTS IN THE PARAMETER STATEMENT.
C      2. CHANGE THE INPUT FILE NAME TO REFLECT THE DATA FILE.
C      3. SET  $XA(I,1)$ ,  $XA(I,2)$ , AND  $XA(I,3)$  EQUAL TO  $X(I)$ ,  $X(I)**2$ ,
C      AND  $X(I)**3$ , RESPECTIVELY.
C      4. SET  $XA(I,4)=Y(I)$ ,  $XA(I,5)=WY(I)=CONSTANT/VAR(Y(I))$ ,
C      AND  $XA(I,6)=WX(I)=CONSTANT/VAR(X(I))$ 
C      5. SET THE VALUES FOR THE ABSOLUTE AND RELATIVE ERRORS
C      USED FOR CHECKING CONVERGENCE OF THE NEW  $X(I)$  VALUES
C      AND THE POLYNOMIAL PARAMETER VALUES.
C
C      INTEGER      IDEP, LDB, LDCOV, LDR, LDSCPE, LDX, NCOEF,
C      &             NCOL, NDEP, NOBS, NROOT
C      & PARAMETER (IDEP=1, LDSCPE=1, NCOEF=4, NCOL=6, NDEP=1,
C      &             NOBS=68, NROOT=1, LDB=NCOEF, LDCOV=NCOEF,
C      &             LDR=NCOEF, LDX=NOBS)
C      REAL*8
C      & BF(LDB,NDEP), BI(LDB,NDEP), BO(LDB,NDEP), BB0, BB1,
C      & BB2, BB3, COVBF(LDCOV,NCOEF), COVBO(LDCOV,NCOEF),
C      & D(NCOEF), DFE, DMACH, EPS, ERRABS, ERRL(4), ERRL,
C      & ETA, F, R(LDR,NCOEF), S2O, S2F, SCPE(LDSCPE,NDEP),
C      & SX, SY, TOL, WY, WX, X, XA(LDX,NCOL), XF(LDX,NCOL),
C      & XI(LDX,NCOL), XMAX(NCOEF), XMIN(NCOEF), Y, ZI(NROOT),
C      & ZF(NROOT)
C      INTEGER      I, IDO, IEEERR, IFRQ, IIND, INDDEP(1), INDIND(1),
C      &             INFO(NROOT), INTCEP, IRANK, ISUB, ITERNLS, IWT,
C      &             ITMAX, K, NRMIS, NROW
C      CHARACTER    CLABEL1(2)*5, CLABEL2(5)*7, FMT*8, RLABEL(4)*7,
C      &             TITLE1*23, TITLE2*26
C      EXTERNAL     DMACH, DRCOV, DRGIVN, DWRRRL, DZREAL, F, MYHANDLER
C      COMMON       BB0, BB1, BB2, BB3, WX, WY, X, Y
C
C      DATA ASSIGNMENTS FOR DWRRRL
C
C      DATA RLABEL /'BETA(0)', 'BETA(1)', 'BETA(2)', 'BETA(3)'/
C      CLABEL1(1)= 'NONE'
C      CLABEL1(2)= 'NONE'
C      DATA CLABEL2 /' ', 'BETA(0)', 'BETA(1)', 'BETA(2)', 'BETA(3)'/
C      DATA FMT/'(D16.8)'/
C      DATA TITLE1 /'REGRESSION COEFFICIENTS'/
C      DATA TITLE2 /'VARIANCE-COVARIANCE MATRIX'/
C
C      PARAMETER VALUES FOR DRGIVN AND DRCOV
C
C      IDO=0
C      NROW=NOBS
C      INTCEP=1
C      IIND=-3
C      INDIND(1)=0
C      INDDEP(1)=4
C      IFRQ=0
C      IWT=5
C      ISUB=1
C      TOL=100*DMACH(4)
C
C      REGRESSION DATA INPUT
C
C      OPEN (9,FILE='calib1.txt',STATUS='OLD')
C      REWIND 9
C      DO 50 I=1,NOBS
C      READ(9,*) I, Y, SY, X, SX
C      XA(I,1)=X

```

```

XA(I,2)=X**2
XA(I,3)=X**3
XA(I,4)=Y
XA(I,5)=1.0D0
XA(I,6)=1.0D0
C
C      ASSIGN THE MEASURED DATA AS INITIAL VALUES
C
XI(I,1)=X
XI(I,2)=X**2
XI(I,3)=X**3
XI(I,4)=Y
XI(I,5)=1.0D0
XI(I,6)=1.0D0
50 CONTINUE
CLOSE(9)
C
C      PERFORM A STANDARD L.S. ANALYSIS OF THE DATA.
C
CALL DRGIVN (IDO, NROW, NCOL, XA, LDX, INTCEP, IIND,
&          INDIND, IDEP, INDDEP, IFRQ, IWT, ISUB, TOL,
&          BO, LDB, R, LDR, D, IRANK, DFE, SCPE, LDSCPE,
&          NRMIS, XMIN, XMAX)
S20=SCPE(1,1)/DFE
CALL DRCOV (NCOEF, R, LDR, S20, COVBO, LDCOV)
C
C      PERFORM A NON-STANDARD LEAST SQUARES ANALYSIS OF THE DATA
C
C      ASSIGN THE STANDARD L.S. PARAMETER ESTIMATES AS INITIAL VALUES
C
BI(1,1)=BO(1,1)
BI(2,1)=BO(2,1)
BI(3,1)=BO(3,1)
BI(4,1)=BO(4,1)
DO 65 I=1,NOBS
XF(I,4)=XA(I,4)
XF(I,5)=XA(I,5)
XF(I,6)=XA(I,6)
65 CONTINUE
C
C      SET THE PARAMETER VALUES FOR DZREAL
C
ERRABS=1.0D-07
ERRREL=1.0D-07
EPS=1.0D-05
ETA=1.0D-02
ITMAX=100
C
C      CALCULATE THE ROOT OF d(i)
C      d(i)=D/DX OF THE WEIGHTED SUM OF THE SQUARED ERRORS
C
ITERNSLS=1
75 DO 100 K=1,NOBS
ZI(1)=XI(K,1)
Y=XA(K,4)
X=XA(K,1)
WY=XA(K,5)
WX=XA(K,6)
BB0=BI(1,1)
BB1=BI(2,1)
BB2=BI(3,1)
BB3=BI(4,1)
CALL DZREAL(F, ERRABS, ERRREL, EPS, ETA, NROOT, ITMAX, ZI, ZF,
&          INFO)
XF(K,1)=ZF(1)
XF(K,2)=ZF(1)**2
XF(K,3)=ZF(1)**3
100 CONTINUE
C
C      PERFORM A LEAST SQUARES ANALYSIS USING THE NEW VALUES FOR THE
C      INDEPENDENT VARIABLES X AND CHECK TO SEE IF THE PARAMETER VALUES
C      HAVE CONVERGED
C
CALL DRGIVN (IDO, NROW, NCOL, XF, LDX, INTCEP, IIND,
&          INDIND, IDEP, INDDEP, IFRQ, IWT, ISUB, TOL,
&          BF, LDB, R, LDR, D, IRANK, DFE, SCPE, LDSCPE,
&          NRMIS, XMIN, XMAX)
ITERNSLS=ITERNSLS+1

```

```

C      CHECK THE NEW PARAMETER VALUES FOR CONVERGENCE
C
ERRB(1)=ABS((BF(1,1)-BI(1,1))/BI(1,1))
ERRB(2)=ABS((BF(2,1)-BI(2,1))/BI(2,1))
ERRB(3)=ABS((BF(3,1)-BI(3,1))/BI(3,1))
ERRB(4)=ABS((BF(4,1)-BI(4,1))/BI(4,1))
IF (ERRB(1).GT.ERRABS .OR. ERRB(2).GT.ERRABS .OR.
& ERRB(3).GT.ERRABS .OR. ERRB(4).GT.ERRABS) THEN
    BI(1,1)=BF(1,1)
    BI(2,1)=BF(2,1)
    BI(3,1)=BF(3,1)
    BI(4,1)=BF(4,1)
C
    DO 150 I=1,NOBS
    XI(I,1)=XF(I,1)
150    CONTINUE
C
    IF (ITERNSLS.GT.1000) GO TO 500
    GO TO 75
ELSE
    SX=0.0D0
    SY=0.0D0
    DO 175 I=1,NOBS
    SX=SX+XA(I,6)*(XA(I,1)-XF(I,1))**2
    SY=SY+XA(I,5)*(XA(I,4)-(BF(1,1)+BF(2,1)*XF(I,1)+BF(3,1)*XF(I,2)
& +BF(4,1)*XF(I,3)))**2
175 CONTINUE
    S2F=SCPE(1,1)/DFE
    CALL DRCOV(B, NCOEF, R, LDR, S2F, COVBF, LDCOV(B))
    END IF
C
C      OUTPUT THE RESULTS
C
WRITE (6,195)
WRITE (6,200)
CALL DWRRRL (TITLE1, LDB, NDEP, BO, LDB, 0, FMT, RLABEL, CLABEL1)
WRITE (6,250)
CALL DWRRRL (TITLE2, LDCOV(B), NCOEF, COVBO, LDCOV(B), 0,
& FMT, RLABEL, CLABEL2)
& WRITE (6,300) S2F
WRITE (6,350)
CALL DWRRRL (TITLE1, LDB, NDEP, BF, LDB, 0, FMT, RLABEL, CLABEL1)
WRITE (6,250)
CALL DWRRRL (TITLE2, LDCOV(B), NCOEF, COVBF, LDCOV(B), 0,
& FMT, RLABEL, CLABEL2)
& WRITE (6,400) S2F
WRITE (6,410) SX, SY, SX+SY
WRITE (6,450) ITERNSLS
C
195 FORMAT (/, 5X, 'REGRESSION ANALYSIS OF (a/W) -VS- N.DCEP', //)
200 FORMAT (/, 5X, 'STANDARD L.S. CURVE FIT RESULTS', /)
250 FORMAT (/)
300 FORMAT (/,5X, ' VARIANCE ESTIMATE FOR THE CURVE FIT = ',
& D16.10)
350 FORMAT (///, 5X, 'ERROR-IN-VARIABLES L.S. CURVE FIT RESULTS',
& /)
400 FORMAT (/,5X, ' VARIANCE ESTIMATE FOR THE CURVE FIT = ',
& D16.10)
410 FORMAT (/,2X, 'SX= ', D16.10, 2X, 'SY= ', D16.10, 2X,
& 'SX+SY= ', D16.10)
450 FORMAT (///, 5X, 'NUMBER OF ITERATIONS =',I3)
GO TO 600
C
500 WRITE (6,550) ITERNSLS
550 FORMAT (///, 5X, '***** CONVERGENCE PROBLEMS! *****')
C
600 END
C
C      FUNCTION DEFINITION FOR THE DERIVATIVE OF THE WEIGHTED
C      SUM OF THE ERRORS SQUARED W.R.T. THE INDEPENDENT VARIABLE
C
REAL*8 FUNCTION F(Z)
REAL*8 BB0, BB1, BB2, BB3, WY, WX, X, Y, Z
COMMON BB0, BB1, BB2, BB3, WX, WY, X, Y
F=-WY*(Y-BB0-BB1*Z-BB2*Z**2-BB3*Z**3)*
& (BB1+2*BB2*Z+3*BB3*Z**2)-WX*(X-Z)
RETURN
END

```

PROGRAM CAL2

MODIFIED FOR DCEP CALIBRATION ANALYSIS; 16-JUL-95.
2ND ORDER POLYNOMIAL FIT OF: $(\kappa Bv/I)(1-a/W)/(a/W) - vs. - (a/W)$
WITH ALL WEIGHTING FACTORS SET EQUAL TO 1.0.

PROGRAM FOR ERROR-IN-VARIABLES, MULTI-LINEAR LEAST
SQUARES CURVE FITTING USING IMSL ROUTINES.

BOTH THE DEPENDENT AND INDEPENDENT VARIABLES HAVE SOME
MEASUREMENT ERROR ASSOCIATED WITH THEM. THE ANALYSIS
IS DESCRIBED BY W. H. SOUTHWELL, "THE COMPUTER JOURNAL",
VOL. 19, NO. 1, 1974, PP. 69-73.

INSTRUCTIONS FOR USE: THE PROGRAM CURRENTLY WRITTEN TO FIT
A 2ND ORDER POLYNOMIAL $X=B_0+B_1*Y+B_2*Y**2$ TO THE DATA SET
{Y(I),X(I)}, I=1,...,NOBS, WHERE THE Y(I) AND THE X(I) ARE
SUBJECT TO UNCERTAINTY DUE TO MEASUREMENT ERRORS.
1. SET NOBS=NUMBER OF DATA POINTS IN THE PARAMETER STATEMENT.
2. CHANGE THE INPUT FILE NAME TO REFLECT THE DATA FILE.
3. SET XA(I,1), XA(I,2), AND XA(I,3) EQUAL TO Y(I), Y(I)**2,
AND 0, RESPECTIVELY.
4. SET XA(I,4)=X(I)*(1-Y(I))/Y(I),
XA(I,5)=WY(I)=CONSTANT/VAR(X(I)),
AND XA(I,6)=WX(I)=CONSTANT/VAR(Y(I))
5. SET THE VALUES FOR THE ABSOLUTE AND RELATIVE ERRORS
USED FOR CHECKING CONVERGENCE OF THE NEW Y(I) VALUES
AND THE POLYNOMIAL PARAMETER VALUES.

```

INTEGER      IDEP, LDB, LDCOV, LDR, LDSCPE, LDX, NCOEF,
&            NCOL, NDEP, NOBS, NROOT
PARAMETER (IDEP=1, LDSCPE=1, NCOEF=4, NCOL=6, NDEP=1,
&            NOBS=68, NROOT=1, LDB=NCOEF, LDCOV=NCOEF,
&            LDR=NCOEF, LDX=NOBS)
REAL*8      BF(LDB,NDEP), BI(LDB,NDEP), BO(LDB,NDEP), BB0, BB1,
&            BB2, BB3, COVEF(LDCOV,NCOEF), COVBO(LDCOV,NCOEF),
&            D(NCOEF), DFE, DMACH, EPS, ERRABS, ERRE(4), ERRREL,
&            ETA, F, R(LDR,NCOEF), S2O, S2F, SCPE(LDSCPE,NDEP),
&            SX, SY, TOL, WY, WX, X, XA(LDX,NCOL), XF(LDX,NCOL),
&            XI(LDX,NCOL), XMAX(NCOEF), XMIN(NCOEF), Y, ZI(NROOT),
&            ZF(NROOT)
INTEGER      I, IDO, IEEEERR, IFRQ, IIND, INDDEP(1), INDIND(1),
&            INFO(NROOT), INTCEP, IRANK, ISUB, ITERNSLS, IWT,
&            ITMAX, K, NRMISS, NROW
CHARACTER    CLABEL1(2)*5, CLABEL2(5)*7, FMT*8, RLABEL(4)*7,
&            TITLE1*23, TITLE2*26
EXTERNAL    DMACH, DRCOV, DRGIVN, DWRRRL, DZREAL, F, MYHANDLER
COMMON      BB0, BB1, BB2, BB3, WX, WY, X, Y

```

DATA ASSIGNMENTS FOR DWRRRL

```

DATA RLABEL /'BETA(0)', 'BETA(1)', 'BETA(2)', 'BETA(3)'/
CLABEL1(1)= 'NONE'
CLABEL1(2)= 'NONE'
DATA CLABEL2 /' ', 'BETA(0)', 'BETA(1)', 'BETA(2)', 'BETA(3)'/
DATA FMT/'(D16.8)'/
DATA TITLE1 /'REGRESSION COEFFICIENTS'/
DATA TITLE2 /'VARIANCE-COVARIANCE MATRIX'/

```

PARAMETER VALUES FOR DRGIVN AND DRCOV

```

IDO=0
NROW=NOBS
INTCEP=1
IIND=-3
INDIND(1)=0
INDDEP(1)=4
IFRQ=0
IWT=5
ISUB=1
TOL=100*DMACH(4)

```

REGRESSION DATA INPUT

```

OPEN (9,FILE='calib1.txt',STATUS='OLD')
REWIND 9
DO 50 I=1,NOBS
READ(9,*) I, Y, SY, X, SX

```

```

      XA(I,1)=Y
      XA(I,2)=Y**2
      XA(I,3)=0.0D0
      XA(I,4)=X
      XA(I,5)=1.0D0
      XA(I,6)=1.0D0
C
C      ASSIGN THE MEASURED DATA AS INITIAL VALUES
C
      XI(I,1)=Y
      XI(I,2)=Y**2
      XI(I,3)=0.0D0
      XI(I,4)=X
      XI(I,5)=1.0D0
      XI(I,6)=1.0D0
50 CONTINUE
      CLOSE(9)
C
C      PERFORM A STANDARD L.S. ANALYSIS OF THE DATA.
C
      CALL DRGIVN (IDO, NROW, NCOL, XA, LDX, INTCEP, IIND,
&                INDIND, IDEP, INDDEP, IFRQ, IWT, ISUB, TOL,
&                BO, LDB, R, LDR, D, IRANK, DFE, SCPE, LDSCPE,
&                NRMISS, XMIN, XMAX)
      S20=SCPE(1,1)/DFE
      CALL DRCOVN (NCOEF, R, LDR, S20, COVBO, LDCOVN)
C
C      PERFORM A NON-STANDARD LEAST SQUARES ANALYSIS OF THE DATA
C
C      ASSIGN THE STANDARD L.S. PARAMETER ESTIMATES AS INITIAL VALUES
C
      BI(1,1)=BO(1,1)
      BI(2,1)=BO(2,1)
      BI(3,1)=BO(3,1)
      BI(4,1)=BO(4,1)
      DO 65 I=1,NOBS
      XF(I,4)=XA(I,4)
      XF(I,5)=XA(I,5)
      XF(I,6)=XA(I,6)
65 CONTINUE
C
C      SET THE PARAMETER VALUES FOR DZREAL
C
      ERRABS=1.0D-07
      ERRREL=1.0D-07
      EPS=1.0D-05
      ETA=1.0D-02
      ITMAX=100
C
C      CALCULATE THE ROOT OF d(i)
C      d(i)=D/DX OF THE WEIGHTED SUM OF THE SQUARED ERRORS
C
      ITERNLS=1
75 DO 100 K=1,NOBS
      ZI(1)=XI(K,1)
      Y=XA(K,4)
      X=XA(K,1)
      WY=XA(K,5)
      WX=XA(K,6)
      BB0=BI(1,1)
      BB1=BI(2,1)
      BB2=BI(3,1)
      BB3=BI(4,1)
      CALL DZREAL(F, ERRABS, ERRREL, EPS, ETA, NROOT, ITMAX, ZI, ZF,
&                INFO)
      XF(K,1)=ZF(1)
      XF(K,2)=ZF(1)**2
      XF(K,3)=0
100 CONTINUE
C
C      PERFORM A LEAST SQUARES ANALYSIS USING THE NEW VALUES FOR THE
C      INDEPENDENT VARIABLES X AND CHECK TO SEE IF THE PARAMETER VALUES
C      HAVE CONVERGED
C
      CALL DRGIVN (IDO, NROW, NCOL, XF, LDX, INTCEP, IIND,
&                INDIND, IDEP, INDDEP, IFRQ, IWT, ISUB, TOL,
&                BF, LDB, R, LDR, D, IRANK, DFE, SCPE, LDSCPE,
&                NRMISS, XMIN, XMAX)

```

```

      ITERNLSL=ITERNSLS+1
C
C      CHECK THE NEW PARAMETER VALUES FOR CONVERGENCE
C
      ERRB(1)=ABS((BF(1,1)-BI(1,1))/BI(1,1))
      ERRB(2)=ABS((BF(2,1)-BI(2,1))/BI(2,1))
      ERRB(3)=ABS((BF(3,1)-BI(3,1))/BI(3,1))
      ERRB(4)=ABS((BF(4,1)-BI(4,1))/BI(4,1))
C
      IF (ERRB(1).GT.ERRABS .OR. ERRB(2).GT.ERRABS .OR.
&      ERRB(3).GT.ERRABS .OR. ERRB(4).GT.ERRABS) THEN
        BI(1,1)=BF(1,1)
        BI(2,1)=BF(2,1)
        BI(3,1)=BF(3,1)
        BI(4,1)=BF(4,1)
C
        DO 150 I=1,NOBS
          XI(I,1)=XF(I,1)
150      CONTINUE
C
        IF (ITERNSLS.GT.1000) GO TO 500
        GO TO 75
      ELSE
        SX=0.0D0
        SY=0.0D0
        DO 175 I=1,NOBS
          SX=SX+XA(I,6)*(XA(I,1)-XF(I,1))**2
          SY=SY+XA(I,5)*(XA(I,4)-(BF(1,1)+BF(2,1)*XF(I,1)+BF(3,1)*XF(I,2)
&          +BF(4,1)*XF(I,3)))**2
175      CONTINUE
        S2F=SCPE(1,1)/DFE
        CALL DRCOV (NCOEF, R, LDR, S2F, COVBF, LDCOV)
        END IF
C
C      OUTPUT THE RESULTS
C
      WRITE (6,195)
      WRITE (6,200)
      CALL DWRRRL (TITLE1, LDB, NDEP, BO, LDB, 0, FMT, RLABEL, CLABEL1)
      WRITE (6,250)
      CALL DWRRRL (TITLE2, LDCOV, NCOEF, COVBO, LDCOV, 0,
&      FMT, RLABEL, CLABEL2)
      WRITE (6,300) S20
      WRITE (6,350)
      CALL DWRRRL (TITLE1, LDB, NDEP, BF, LDB, 0, FMT, RLABEL, CLABEL1)
      WRITE (6,250)
      CALL DWRRRL (TITLE2, LDCOV, NCOEF, COVBF, LDCOV, 0,
&      FMT, RLABEL, CLABEL2)
      WRITE (6,400) S2F
      WRITE (6,410) SX, SY, SX+SY
      WRITE (6,450) ITERNSL
C
195  FORMAT (/, 5X, 'REGRESSION ANALYSIS OF N.DCEP -VS- (a/W)', //)
200  FORMAT (/, 5X, 'STANDARD L.S. CURVE FIT RESULTS', //)
250  FORMAT (//)
300  FORMAT (/,5X, ' VARIANCE ESTIMATE FOR THE CURVE FIT = ',
&      D16.10)
350  FORMAT (///, 5X, 'ERROR-IN-VARIABLES L.S. CURVE FIT RESULTS',
&      //)
400  FORMAT (/,5X, ' VARIANCE ESTIMATE FOR THE CURVE FIT = ',
&      D16.10)
410  FORMAT (/,2X, 'SX= ', D16.10, 2X, 'SY= ', D16.10, 2X,
&      'SX+SY= ', D16.10)
450  FORMAT (///, 5X, 'NUMBER OF ITERATIONS =',I3)
      GO TO 600
500  WRITE (6,550) ITERNSL
550  FORMAT (///, 5X, '***** CONVERGENCE PROBLEMS! *****')
600  END
C
C      FUNCTION DEFINITION FOR THE DERIVATIVE OF THE WEIGHTED
C      SUM OF THE ERRORS SQUARED W.R.T. THE INDEPENDENT VARIABLE
C
      REAL*8 FUNCTION F(Z)
      REAL*8 BB0, BB1, BB2, BB3, WY, WX, X, Y, Z
      COMMON BB0, BB1, BB2, BB3, WX, WY, X, Y
      F=-WY*(Y-BB0-BB1*Z-BB2*Z**2-BB3*Z**3)*
&      (BB1+2*BB2*Z+3*BB3*Z**2)-WX*(X-Z)
      RETURN
      END

```

Grid Level Parallelization of an Implicit Solution of the 3D Navier-Stokes Equations

Karen A. Tomko
Assistant Professor
Department of Computer Science and Engineering

Wright State University
Colonel Glen Highway
Dayton, OH 45435-0001

Final Report for:
Summer Faculty Research Program
Wright Laboratory
Aeromechanics Division

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and
Wright Laboratory

September 1996

Grid Level Parallelization of an Implicit Solution of the 3D Navier-Stokes Equations

Karen A. Tomko

Assistant Professor

Department of Computer Science and Engineering
Wright State University

Abstract

A parallel version of the FDL3DI application from Wright Laboratory was developed using MPI on a Cray T3D. This report describes the parallel Chimera version of FDL3DI which solves the three-dimensional Navier-Stokes equations for multiple overlapped grids. A simple approach to parallelizing the Chimera method is taken. Each grid is assigned to a separate processor and the interpolated boundary points are exchanged between processors as necessary. This approach requires few modifications to the source and relatively little communication between processors.

The execution time for an 8 grid flow problem was only 1.6 times greater on eight processors of the T3D than on a single processor of the Cray C916 vector processor. These preliminary results on the T3D demonstrate that parallel systems are competitive with the Cray vector systems for this application.

Grid Level Parallelization of an Implicit Solution of the 3D Navier-Stokes Equations

Karen A. Tomko

1 Introduction

The potential of massively parallel computer systems (MPPs) for the solution of large computational aeroscience problems has yet to be realized. In order for an application to attain high performance on a parallel computer it is essential that the problem is decomposed into subproblems such that the work is shared among the processors. This decomposition into subproblems, *domain decomposition*, must be done in such a way that the work load is distributed equitably among the processors, communication cost between the processors is kept as small as possible and the numerical accuracy of the problem is not compromised. However, the goals of balancing the work load, minimizing communication and maintaining accuracy within an application are often at odds, complicating the task.

During my summer faculty position with the Aeromechanics Division of Wright Laboratory I implemented a parallel version of the FDL3DI application. FDL3DI solves the three-dimensional Navier-Stokes equations using the approximate-factorization algorithm of Beam and Warming in conjunction with a newton subiteration procedure to enhance the accuracy for rapid fluid motion. I parallelized the Chimera version of FDL3DI which solves the Navier-Stokes equations for multiple overlapped grids.

The Chimera method can be parallelized by simply assigning each grid to a separate processor and communicating the interpolation points as necessary. This approach has two main strengths. Firstly, it is conceptually very simple and requires relatively little modification to the source code. Secondly, since only the interpolation points due to the overlap between grids must be communicated between processors, the amount of communication is kept small.

The parallel version of FDL3DI has been implemented using the MPI message passing library and tested on a Cray T3D parallel system. The correctness of the parallel program was validated by solving a temporally evolving Couette flow problem and comparing the program results with the analytic solution. An eight grid problem on 8 processors of the T3D took only 1.6 times as long as the same problem on a single processor of the Cray C916 vector computer. This demonstrates that parallel computer such as the T3D are a viable alternative to the Cray vector computer for this application.

There are two important issues which complicate this parallel approach for FDL3DI which must still be addressed: 1) load balancing - subgrids may vary greatly in size so the work may not be distributed equitably across the processors 2) numerical accuracy - splitting grids to achieve better load balancing may adversely effect the accuracy of the simulation.

This report is organized as follows. Related research is discussed in Section 2. The original and parallel versions of the FDL3DI application are described in Sections 3 and 4 respectively. Experimental runs on the T3D are described in Section 5 followed by a discussion of the single processor application performance in Section 6. We conclude with some suggested next steps to be taken toward achieving a

scalable parallel FDL3DI.

2 Related Work

Many research and industrial groups have parallelized their CFD applications using different techniques and parallel machines. Several of these groups have used a similar approach to ours for their multiple grid applications [1, 6, 3, 16]. Alonso *et al.* and Krishnan demonstrate how load imbalance due to grid size variance limits the scalability of their solution as noted above.

As also noted, numerical accuracy must be maintained when grids are broken into subgrids for parallelization. Meakin [8] and Lutton and Visbal [7] discuss the numerical accuracy of Chimera (or overset) grid solutions for unsteady flows. Both agree that using Chimera grid methods can compute flow fields with acceptable accuracy. However, the step size may have to be reduced or the number of newton subiterations increased as the number of subgrids is increased.

An alternative approach is to parallelize the flow solver, which uses an *alternating-direction-implicit* (ADI) direct solution method. A multi-partition method for ADI solvers has been used in POVERFLOW, a parallel version of the NASA Overflow code [4, 12, 11]. A uni-partition method is used in [10] by Ryan and Weeratunga. Either of these methods require a great deal more programming effort and have higher levels of communication than our approach. Their advantage is that they do not suffer from load imbalance or numerical inaccuracy. One of these methods could be used in conjunction with our method to provide a hybrid parallelization scheme. Such a hybrid approach is outlined in [14]

Many aerospace applications which are similar to FDL3DI are implemented as unstructured mesh applications. A good summary of the issues involved and common solutions for parallelizing these applications is given by Barth in [2]. These methods are not directly applicable to FDL3DI since it uses structured meshes.

Most current implementations of parallel applications use explicit message passing either with MPI or PVM as I have done. However, data parallel implementations of the Navier-Stokes equations are discussed in [5] and [9]. Authors of these articles conclude that current implementations are not yet robust enough to support large complex applications such as FDL3DI. As High Performance Fortran, a data parallel Fortran, matures and becomes more widely supported, it may become desirable to reconsider this decision.

3 Description of FDL3DI with Chimera Support

The FDL3DI application is used by several researchers at Wright Laboratory and elsewhere and has been modified and enhanced by many of them since it was initially developed in 1988. In this report I describe a version of the application which supports Chimera (or overset) grids. References to the FDL3DI application throughout this document are to this version of the code.

The FDL3DI application consists of approximately 9000 lines of Fortran code. A complete list of the subroutines in FDL3DI and a call graph showing the hierarchical relationships between subroutines are given in Appendix A. A high level sketch of the algorithm is given below.

Algorithm 1 (FDL3DI) *Calculate the three-dimensional compressible Navier-Stokes equations with the implicit Beam-Warming algorithm using the Chimera method for overlapped grids.*

1. *Read in the program parameters file and the restart file*
2. *Increment the time step*
3. *For each grid*
 - (a) *Perform initialization*
 - (b) *Solve the flow equations for the current grid*
 - (c) *Calculate the flow values for the Chimera interpolated boundary points*
4. *For each grid*
 - (a) *Update Chimera boundaries with data from neighboring grids*
 - (b) *Apply physical boundary conditions*
5. *Repeat Steps 3 and 4 for each newton subiteration*
6. *Repeat Steps 2 thru 5 for each timestep*
7. *Output results to a new restart file*

Within each timestep of the FDL3DI simulation, we perform n newton subiterations. Within each subiteration the flow equations are solved independently for each grid then the Chimera interpolated boundary values, which the grid supplies to neighboring grids, are computed. After all grids have been updated, each grid is visited again and the boundary conditions are applied for the Chimera grid boundaries and for the physical boundaries.

4 Parallelization

A simple approach has been taken for parallelizing FDL3DI. Each input grid is assigned to a separate processor. The flow equations for each grid are solved independently in parallel and the interpolated boundary values are also updated in parallel. The boundary data is exchanged between processors then, on each processor, the Chimera boundary conditions and the physical boundary conditions are applied to the assigned grid.

The Single Program Multiple Data (SPMD) parallel programming style is used. The code running on each processor is identical and the processor identification number is used to determine which grid is assigned to each processor. The MPI message passing library is used for interprocessor communication. Point-to-point communication using send and receive calls are used to exchange the Chimera boundary data between processors.

The parallel algorithm is sketched out below.

Algorithm 2 (Parallel FDL3DI) *Calculate, in parallel, the three-dimensional compressible Navier-Stokes equations with the implicit Beam-Warming algorithm using the Chimera method for overlapped grids. This algorithm runs simultaneously on each processor. Grids are assigned to processors by a mapping between processor id numbers and grid numbers.*

1. Initialize message passing system
2. Read in the program parameters file and the restart file (the entire restart file is read in redundantly on each processor.)
3. Determine which boundary data elements must be exchanged between processors
4. Increment the time step
5. For the grid assigned to this processor
 - (a) Perform initialization
 - (b) Solve flow equation for current grid
 - (c) Calculate flow values for Chimera boundary points
 - (d) Send flow values for Chimera boundary points to each processor requiring the data
 - (e) Wait until all interpolated boundary data are received from other processors
 - (f) Update Chimera boundaries with data from neighboring grids
 - (g) Apply physical boundary conditions
6. Repeat Step 5 for each newton subiteration
7. Repeat Steps 4 thru 6 for each timestep
8. Send final flow results for each grid to processor 0 for output
9. Output results to a new restart file

To implement Algorithm 2 changes were made to three of the FDL3DI subroutines and 11 additional routines were developed to perform the message passing as listed in Table 1.

5 Experimental Results

All experimental runs were executed, in batch mode, on the 128 processor Cray T3D at the Ohio Supercomputer Center. To run FDL3DI you must have a restart file containing the grid geometry and initial conditions. A small Fortran program, GENINPUT, was used to generate the simple grids used for our test cases and the PEGSUS program [13] was used to determine the Chimera boundaries between grids. Both GENINPUT and PEGSUS had to be recompiled and executed on the Cray T3D parallel system so that they would produce input files for FDL3DI in the appropriate format. In addition, some utilities were used to convert the unformatted output files to a single precision SGI data format for use with the

Modified	New
chd3d()	qb_send_setup(), qb_recv_setup()
intrbc()	send_qb(), recv_qb()
updibc()	wait_qb_send(), wait_qb_recv()
	flv_setup()
	send_flv(), recv_flv()
	wait_flv_send(), wait_flv_recv()

Table 1: New/Modified code in Parallel FDL3DI

data analysis and visualization packages. The ITRANS program for converting from Cray vector format to SGI format was modified to convert Cray T3D formatted data to SGI format.

Two flow problems were used for debugging and verification of Parallel FDL3DI, a uniform flow example and a temporally evolving Couette flow example. The initial uniform flow example used a uniformly spaced grid of $81 \times 41 \times 41$ which was split into two equal sized grids of $46 \times 41 \times 41$ with an overlap of 11 grid points. The application was executed for 100 iterations using 1 newton subiteration per timestep on both the Cray C90 and on 1 and 2 processors of the Cray T3D. Results are given in Table 2. A comparison of the single grid runs on both systems shows that a single processor of the T3D is about 7 times slower than a single processor of the C90. The parallel speedup for a two processor run on the T3D is 1.76.

An eight grid example was also used to more thoroughly test the message passing code. Each grid is uniformly spaced and has $23 \times 23 \times 23$ grid points with a 5 point overlap at Chimera boundaries. The eight grids together form a cube with two grids in each dimension. In the middle of the cube all of the grids overlap, hence each grid shares a boundary with every other grid. Consequently, a data exchange between every pair of processors is required for each newton subiteration of FDL3DI.

Two examples were run on the 8 grid cube, a uniform flow and a temporally evolving Couette flow. Couette flow is a classic problem in fluid dynamics, it was chosen because the Navier-Stokes equations reduce to a single linear partial differential equation, thus the correctness of the solution can easily be confirmed.

Here is a brief description of the physical problem for those unfamiliar with Couette flow. Consider two flat plates which are separated by a distance h . The plates are infinite in two directions, and the space between the plates is entirely filled with an incompressible fluid. Both plates are initially stationary, but at time $t = 0.0$ the upper plate is instantaneously accelerated to a constant velocity, U_{ref} . Since the plates and fluid are infinite in two directions, the properties of the fluid at any point at any given time depends spatially only upon y , which is the distance from the fixed lower plate.

Figure 1 gives velocity vs. distance curves for the test example, where y is the vertical distance normalized by h , and U is the velocity in the direction of the moving plate motion (in this example taken as x) normalized by U_{ref} . Also, t is the nondimensional time given by $t = t^* \times U_{ref}/h$ where t^* is the actual physical time. For the results in Figure 1, the Reynolds number, Re was taken as equal to 1.0^1 .

¹A single parameter, Re - the Reynolds number, appears in the flow equations. Reynolds number is a nondimensional parameter which is a ratio of the inertial to viscous forces acting upon the fluid.

Machine	Grids	Processors	User Time (seconds)
Cray C90	1	1	385
Cray T3D	1	1	2683
	2	1	3258
	2	2	1521

Table 2: Execution Times for Uniform Flow

Machine	Flow	Grids	Processors	User Time (seconds)
Cray C90	Uniform	8	1	410
	Couette	8	1	548
Cray T3D	Uniform	8	8	675
	Couette	8	8	860

Table 3: Execution Times for Couette Flow

Figure 1 shows a comparison between the analytic solution to the Couette flow problem and the solution of parallel FDL3DI. For each value of t , both the analytic and FDL3DI curves are shown. For $t = 0.01$ thru $t = 0.04$ the curves are in exact agreement. For the steady state solution at time $t = .44$, there is an expected discrepancy with the analytic solution. In order to reach the steady state quickly with FDL3DI, the number of newton subiterations per timestep was reduced from 3 to 1 and the timestep size was increase causing some inaccuracy in the solution. However the parallel results matched the results of the Cray vector version of the application when the same input parameters are used.

For the 8 grid uniform flow case the application was executed for 100 iterations using 3 newton subiteration per timestep on both the Cray C90 and on 8 processors of the Cray T3D. The 8 grid Couette flow case was executed for 100 iterations using 4 newton subiterations per timestep. The execution time results are given in Table 3. An 8 processor run on the T3D is about 1.6 times slower than the same run on the C90 which demonstrates that the T3D is a viable computer platform for this application.

6 Single Processor Performance

A parallel application can not achieve high performance without making efficient use of each individual processor. To do so, the application must be able to efficiently access program data (i.e. make efficient use of cache memory). In addition, the application should try to minimize or avoid expensive operations.

A thorough discussion on optimizing for uniprocessor caches is given in Chapter 10 of [15]. If a data item is in cache or a register and it is reused later in the application, it is desirable for the reuse to occur while the value is still in a register or in the cache. If we reference a data item which has remained in a register or cache the reference is said to exhibit *locality*. There are many techniques which have been developed to transform loops to help increase locality.

I have used two of the techniques, *loop fusion* and *array contraction* in some of the subroutines in FDL3DI. Loop fusion can increase locality by combining loops when they use the same data. Array

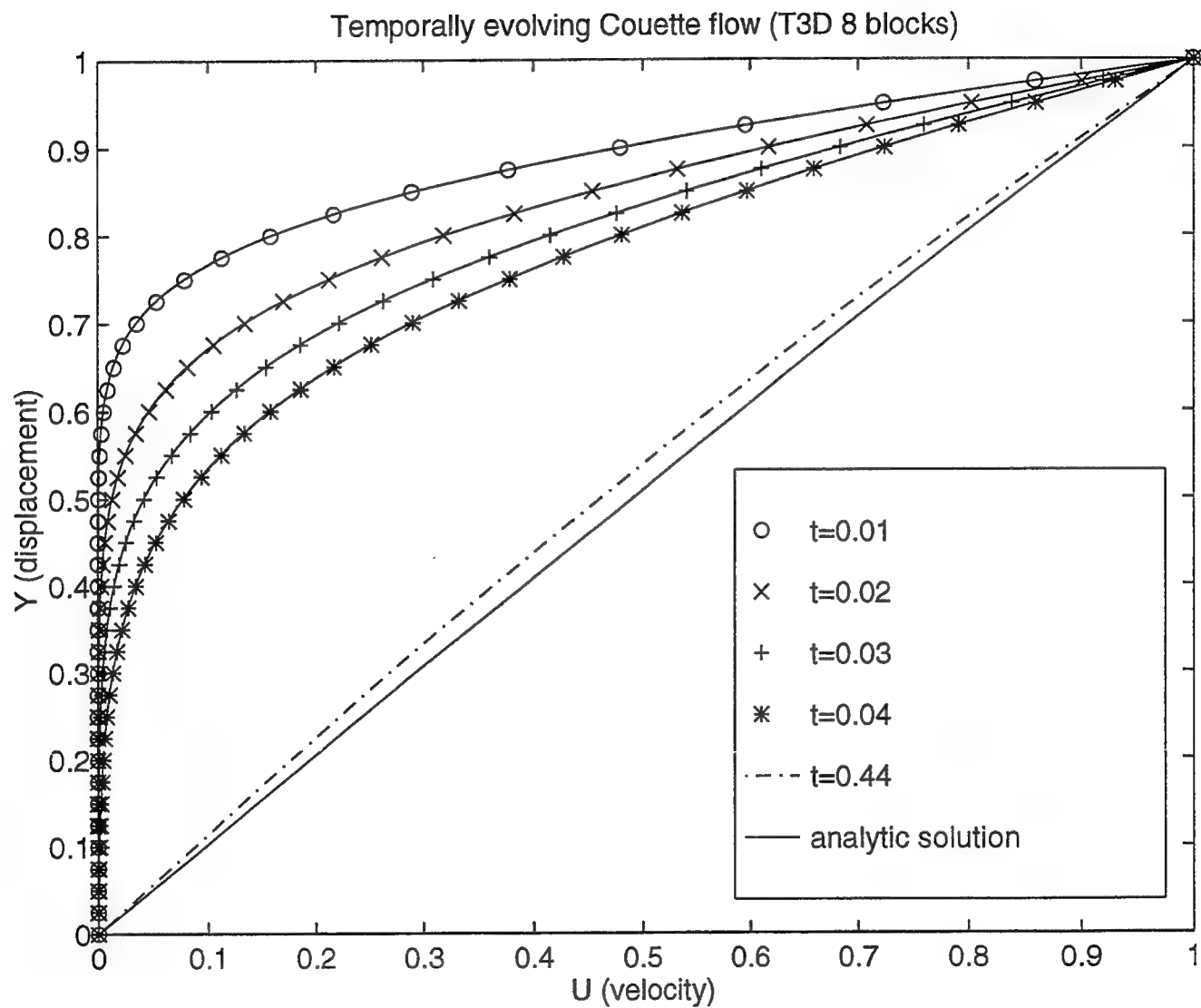


Figure 1: Comparison of Parallel FDL3DI with the analytic solution for the Couette flow example

Subroutine Name	Execution Time (seconds)	
	Original	Optimized
RHSXZ	386.5	283.3
VCROSY	268.4	147.0
VCROSY	156.3	91.0
VCROSY	154.4	90.3
Total	965.6	611.5

Table 4: Execution Times for optimized subroutines of FDL3DI

contraction is the replacement of a temporary array variable (within a loop) with a scalar temporary. This replacement can be done if the array is not used after the loop.

Figure 2 shows how loop fusion and array contraction were used to optimize two loops in the routine VCROSY. Four routines in FDL3DI were optimized in this way. Execution times for these routines were reduced by a total of 36% and VCROSY improved by 45%. The execution times for the original subroutines and for the optimized subroutines are given in Table 4. The times were measured using the Apprentice performance analysis tool on the Cray T3D.

For most multiprocessors floating point divide and square root operations are expensive. Floating point division can often be approximated by replacing the division by a multiply of the reciprocal of the divisor. The cf77 compiler for the T3D will automatically perform this optimization if the compiler flag `-Wf-onoieedivide` is used. Square root is inherently slow because of the iterative nature of the square root algorithm. In the original FDL3DI application, the SQRT intrinsic routine accounted for 14.6% of the total runtime. In order to reduce the cost of square root I stored square root results for later use instead of recalculating the value. Eight subroutines in all were modified and their runtimes before and after the modifications are given in Table 5. The runtimes for most of the routines decreased because they are no longer performing the square root operation. However the runtime of a few of the routines increased, most notably TXIRHS. This is due to the fact that the square root value, which was previously in a scalar temporary, is saved to an array for reuse in NIRHS and the memory accesses to the array have slowed down the routine. Overall, the optimization was not very successful, it reduced the total time for these routines by only 7%.

The optimizations discussed above gave a disappointing runtime improvement of only 11%. The small impact of these changes is due to the fact that we only optimized a handful of routines none of which dominate the runtime of the application. As shown in the Apprentice Report listing (Appendix A) from the original application, there are many routines which take between 1-8% of the runtime. Each of these routines must be optimized to get a significant overall reduction in run time.

Another, difficulty with performing these transformations is that the numerical results will vary slightly from the original results due to reordering of some of the arithmetic operations. This variance makes it difficult to determine if the application is still correct or whether bugs have been introduced in the optimization process. Thus, the application should be formally verified after applying these transformations.

As shown above, simple transformations such as loop fusion and array contraction can potentially

Original Loops

```

C
C COMPUTE FV(XI), GV(XI) AND HV(XI)
C
DO 2 K=KS,KE
  IF((K.EQ.KS).OR.(K.EQ.KE)) THEN
    JJB=JSP1
    JJE=JEM1
  ELSE
    JJB=JS
    JJE=JE
  ENDIF
DO 21 J=JJB,JJE
  XJACI=1./XJAC(I,J,K)
  XMUXJ=XJACI*(XHU(J,K)+EDDY(I,J,K))
  XKTJ=XJACI*XXI*(XHU(J,K)+PRI+EDDY(I,J,K)*PRTI)
  UXI=U(IP,J,K)-U(IM,J,K)
  VXI=V(IP,J,K)-V(IM,J,K)
  WXI=W(IP,J,K)-W(IM,J,K)
  TXI=GXM*(P(IP,J,K)/RHO(IP,J,K)-P(IM,J,K)/RHO(IM,J,K))
  XPHI=XLBD*(UXI*XIX(I,J,K)+VXI*XIY(I,J,K)
1 +WXI*XIZ(I,J,K))
C
C ---- FORM FV(XI) ----
C
WRK(J,K,2,1)=XMUXJ*(2.*UXI*XIX(I,J,K)+XPHI)
WRK(J,K,3,1)=XMUXJ*(UXI*XIY(I,J,K)+VXI*XIX(I,J,K))
WRK(J,K,4,1)=XMUXJ*(UXI*XIZ(I,J,K)+WXI*XIX(I,J,K))
WRK(J,K,5,1)=U(I,J,K)*WRK(J,K,2,1)+V(I,J,K)*WRK(J,K,3,1)
1 +W(I,J,K)*WRK(J,K,4,1)+XKTJ*TXI*XIX(I,J,K)
C
C ---- FORM GV(XI) ----
C
WRK(J,K,2,2)=WRK(J,K,3,1)
WRK(J,K,3,2)=XMUXJ*(2.*VXI*XIY(I,J,K)+XPHI)
WRK(J,K,4,2)=XMUXJ*(VXI*XIZ(I,J,K)+WXI*XIY(I,J,K))
WRK(J,K,5,2)=U(I,J,K)*WRK(J,K,2,2)+V(I,J,K)*WRK(J,K,3,2)
1 +W(I,J,K)*WRK(J,K,4,2)+XKTJ*TXI*XIY(I,J,K)
C
C ---- FORM HV(XI) ----
C
WRK(J,K,2,3)=WRK(J,K,4,1)
WRK(J,K,3,3)=WRK(J,K,4,2)
WRK(J,K,4,3)=XMUXJ*(2.*WXI*XIZ(I,J,K)+XPHI)
WRK(J,K,5,3)=U(I,J,K)*WRK(J,K,2,3)+V(I,J,K)*WRK(J,K,3,3)
1 +W(I,J,K)*WRK(J,K,4,3)+XKTJ*TXI*XIZ(I,J,K)
C
21 CONTINUE
2 CONTINUE
C
DO 3 L=2,5
C
DO 31 K=KS,KE
  IF((K.EQ.KS).OR.(K.EQ.KE)) THEN
    JJB=JSP1
    JJE=JEM1
  ELSE
    JJB=JS
    JJE=JE
  ENDIF
C
DO 31 J=JJB,JJE
  WRK(J,K,L,4)=ETAX(I,J,K)*WRK(J,K,L,1)+
1 ETAY(I,J,K)*WRK(J,K,L,2)+ETAZ(I,J,K)*WRK(J,K,L,3)
  WRK(J,K,L,5)=ZETAX(I,J,K)*WRK(J,K,L,1)+
1 ZETAY(I,J,K)*WRK(J,K,L,2)+ZETAZ(I,J,K)*WRK(J,K,L,3)
C
31 CONTINUE
3 CONTINUE

```

Optimized Loop

```

C
C COMPUTE FV(XI), GV(XI) AND HV(XI)
C
C Combine loops 2 and 3 and unroll loop 3 - ktomko 7/10/96
DO 2 K=KS,KE
  IF((K.EQ.KS).OR.(K.EQ.KE)) THEN
    JJB=JSP1
    JJE=JEM1
  ELSE
    JJB=JS
    JJE=JE
  ENDIF
DO 21 J=JJB,JJE
  XJACI=1./XJAC(I,J,K)
  XMUXJ=XJACI*(XHU(J,K)+EDDY(I,J,K))
  XKTJ=XJACI*XXI*(XHU(J,K)+PRI+EDDY(I,J,K)*PRTI)
  UXI=U(IP,J,K)-U(IM,J,K)
  VXI=V(IP,J,K)-V(IM,J,K)
  WXI=W(IP,J,K)-W(IM,J,K)
  TXI=GXM*(P(IP,J,K)/RHO(IP,J,K)-P(IM,J,K)/RHO(IM,J,K))
  XPHI=XLBD*(UXI*XIX(I,J,K)+VXI*XIY(I,J,K)
1 +WXI*XIZ(I,J,K))
C
C Replace some of the working arrays with scalar
C temporaries - ktomko
C
C ---- FORM FV(XI) ----
C
WRK21=XMUXJ*(2.*UXI*XIX(I,J,K)+XPHI)
WRK31=XMUXJ*(UXI*XIY(I,J,K)+VXI*XIX(I,J,K))
WRK41=XMUXJ*(UXI*XIZ(I,J,K)+WXI*XIX(I,J,K))
WRK51=U(I,J,K)*WRK21+V(I,J,K)*WRK31
1 +W(I,J,K)*WRK41+XKTJ*TXI*XIX(I,J,K)
C
C ---- FORM GV(XI) ----
C
WRK22=WRK31
WRK32=XMUXJ*(2.*VXI*XIY(I,J,K)+XPHI)
WRK42=XMUXJ*(VXI*XIZ(I,J,K)+WXI*XIY(I,J,K))
WRK52=U(I,J,K)*WRK22+V(I,J,K)*WRK32
1 +W(I,J,K)*WRK42+XKTJ*TXI*XIY(I,J,K)
C
C ---- FORM HV(XI) ----
C
WRK23=WRK41
WRK33=WRK42
WRK43=XMUXJ*(2.*WXI*XIZ(I,J,K)+XPHI)
WRK53=U(I,J,K)*WRK23+V(I,J,K)*WRK33
1 +W(I,J,K)*WRK43+XKTJ*TXI*XIZ(I,J,K)
C
WRK(J,K,2,4)=ETAX(I,J,K)*WRK21+
1 ETAY(I,J,K)*WRK22+ETAZ(I,J,K)*WRK23
WRK(J,K,2,5)=ZETAX(I,J,K)*WRK21+
1 ZETAY(I,J,K)*WRK22+ZETAZ(I,J,K)*WRK23
WRK(J,K,3,4)=ETAX(I,J,K)*WRK31+
1 ETAY(I,J,K)*WRK32+ETAZ(I,J,K)*WRK33
WRK(J,K,3,5)=ZETAX(I,J,K)*WRK31+
1 ZETAY(I,J,K)*WRK32+ZETAZ(I,J,K)*WRK33
WRK(J,K,4,4)=ETAX(I,J,K)*WRK41+
1 ETAY(I,J,K)*WRK42+ETAZ(I,J,K)*WRK43
WRK(J,K,4,5)=ZETAX(I,J,K)*WRK41+
1 ZETAY(I,J,K)*WRK42+ZETAZ(I,J,K)*WRK43
WRK(J,K,5,4)=ETAX(I,J,K)*WRK51+
1 ETAY(I,J,K)*WRK52+ETAZ(I,J,K)*WRK53
WRK(J,K,5,5)=ZETAX(I,J,K)*WRK51+
1 ZETAY(I,J,K)*WRK52+ZETAZ(I,J,K)*WRK53
21 CONTINUE
2 CONTINUE

```

Figure 2: Locality enhancing transformation of VCROSLX loops.

Subroutine Name	Execution Time (seconds)	
	Original	Optimized
NIRHS	76.1	50.1
PIRHS	76.1	78.9
SDIAGX	199.6	163.6
SDIAGY	218.1	188.3
SPECX	46.3	47.2
SPECY	82.1	48.1
TXIIRHS	95.6	159.8
TZETRHS	93.0	86.1
TOTAL	886.9	822.5

Table 5: Execution Times for subroutines with square root calls in FDL3DI

reduce the application run time by 35%. Although labor intensive, it is a straight forward to perform these improvements on the rest of the routines in FDL3DI. Other locality improving transformations should also be investigated.

7 Conclusion and Future Work

In this report we have demonstrated that FDL3DI can be parallelized using a grid based decomposition across processors. The T3D is shown to be a feasible alternative to the Cray C90 Vector processors, with a runtime of 1.6 times the runtime on the vector processor.

However, this is only a preliminary step in achieving a scalable version of FDL3DI which can maintain efficiency on large numbers of processors. Several issues still need to be addressed:

1. Automatic grid decomposition
2. Load balancing of multiple grids of varying dimensionality
3. Parallel performance analysis
4. Numerical accuracy vs. parallelization
5. Performance tuning for locality and message passing

The experimental cases included in this report are limited by our inability to quickly decompose grids into equal sized subgrids. An automatic decomposition package is needed to split the regular grids into any number of subgrids. This tool should allow the user control over the degree to which each dimension is split and the amount of overlap between subgrids. In addition the external boundary conditions of the original grid need to be appropriately applied to subgrid boundaries.

Once individual grids can be decomposed into an arbitrary number of equal sized subgrids, an overall load balancing algorithm can be developed. This algorithm should provide a means of distributing multiple Chimera grids of unequal size equitably among a set of processors. Again, this tool should allow the user some control over the decomposition and perform boundary updates appropriately.

The automatic decomposition tools described above will provide the ability to run a sample flow problem using many different grid decompositions. With this ability we can perform a comprehensive

performance evaluation of the application. This performance evaluation should include a comparison of FDL3DI execution on several different parallel platforms including the IBM SP2, Cray T3D and T3E and the SGI Power Challenge system.

Once the performance of the application is understood, it will be important to evaluate how increasing the parallelism effects the numerical accuracy of the application. The parallel application may have to perform extra calculations to compensate for the lag introduced by the Chimera boundaries. We need to determine how to compensate for the inaccuracies as parallelism is increased, and at what point it is no longer beneficial to increase parallelism.

Finally, both locality optimization and message passing optimization need to be considered. There are some simple transformations which can quickly improve locality and thus performance as described in Section 6. Other locality improving transformations should also be considered. In addition, more advanced features of MPI can be used to reduce the number of messages between processors and hence reduce the overhead of message passing in the application. Such optimizations should not compromise the portability of the application.

While much remains to be done, the potential gains of running FDL3DI in parallel are great. An MPI version of FDL3DI can be easily moved onto new parallel systems, providing the flexibility to chose from the available computer resources. In addition, many parallel systems are less heavily used than the vector systems, hence queue wait time is reduced. Most importantly parallel systems can potentially allow the solution of problems which are too large to be solved on the available vector systems, due to the large amounts of memory on the parallel systems. We have demonstrated that execution times on parallel computers are competitive with vector systems. After the steps outlined above have been taken, performance should be better than the vector systems (and at a 10th the cost).

8 Acknowledgements

Thanks to Don Rizzetta for providing the source code, devising the test cases, and answering numerous questions which made this work possible. Thanks to Steve Scherr for his help with the ITRANS utility. Thanks also to Joseph Shang and Miguel Visbal for giving me the opportunity to work with the FIMC research group.

Computational resources for the work presented here were supported in part by a grant of HPC time from the DOD HPC Shared Resource Center, CEWES, Vicksburg, MS. Additional computer resources were provided by the Ohio Super Computer Center, Columbus OH.

A FDL3DI Source Code Charts

Routines {						
; Name	Class	Type				
CHD3D	External	Program	Calls EXIT	SDAMPZ1	External	Subroutine
BTRID	External	Subroutine		SDAMPZKE	External	Subroutine
CHATA	External	Subroutine		SDIAGX	External	Subroutine
CHATB	External	Subroutine		SDIAGY	External	Subroutine
CHATC	External	Subroutine		SDIAGZ	External	Subroutine
CHATV1	External	Subroutine		SETPARM	External	Subroutine
CHATV2	External	Subroutine		SOURCE	External	Subroutine
CHATV3	External	Subroutine		SPECX	External	Subroutine
CONVRT	External	Subroutine		SPECY	External	Subroutine
CXMU1	External	Subroutine		SPECZ	External	Subroutine
DAMPKE	External	Subroutine		SPENTA	External	Subroutine
DAMPX	External	Subroutine		SPENTA3	External	Subroutine
DAMPY	External	Subroutine		SST	External	Subroutine
DAMPZ	External	Subroutine		STEPETAKE	External	Subroutine
FILLMX	External	Subroutine		STEPXIKE	External	Subroutine
FILLMY	External	Subroutine		STEPZETAKE	External	Subroutine
FILLMZ	External	Subroutine		SWEEPI	External	Subroutine
GETIBK	External	Subroutine		SWEEPJ1	External	Subroutine
GETINT	External	Subroutine		SWEEPJ2	External	Subroutine
GST	External	Subroutine		SWEEPK	External	Subroutine
IEDRV	External	Subroutine		SWEEPKKE	External	Subroutine
INITIA	External	Subroutine		TNSTEP	External	Subroutine
INTRBC	External	Subroutine		TXIIRHS	External	Subroutine
L2NORM	External	Subroutine		TZETRHS	External	Subroutine
METRIC	External	Subroutine		UPDTQO	External	Subroutine
NIRHS	External	Subroutine		VCROX	External	Subroutine
NS3D	External	Subroutine		VCROSY	External	Subroutine
OUTPUT	External	Subroutine		VCROSZ	External	Subroutine
PIRHS	External	Subroutine		CXMU	External	Subroutine
PUTBND	External	Subroutine		TURBFS	External	Subroutine
PUTIBK	External	Subroutine		ABS	Intrinsic	Function
PUTINT	External	Subroutine		AMAX1	Intrinsic	Function
PUTQB	External	Subroutine		AMIN1	Intrinsic	Function
RHSKE	External	Subroutine		ATAN	Intrinsic	Function
RHSSUB	External	Subroutine		COS	Intrinsic	Function
RHSXZ	External	Subroutine		EXP	Intrinsic	Function
RHSY	External	Subroutine		FLOAT	Intrinsic	Function
SDAMPX	External	Subroutine		SIN	Intrinsic	Function
SDAMPXYKE	External	Subroutine		SQRT	Intrinsic	Function
SDAMPY	External	Subroutine		BCDRV	Unknown	Subroutine
SDAMPZ	External	Subroutine				

Subroutine List for FDL3DI generated from FORGEExplorer

Call Subchain... {		Status
; Node		
1:	CHD3D	
2:	INITIA	
3:	TURBFS	
4:	PUTINT	
5:	PUTBND	
6:	PUTIBK	
7:	SETPARM	
8:	SST	
9:	INTRBC	
10:	GETINT	
11:	PUTQB	
12:	IEDRV	
13:	INITIA	<See 2>
14:	SETPARM	
15:	GETIBK	
16:	GST	
17:	NS3D	
18:	METRIC	
19:	THSTEP	
20:	THSTEP	
21:	UPDTQO	
22:	THSTEP	
23:	SWEEPJ1	
24:	CXMU	
25:	RHSXZ	
26:	VCROSSY	
27:	CONVRT	
28:	DAMPZ	
29:	SDAMPZ	
30:	SPECZ	
31:	RHSSUB	
32:	SWEEPI	
33:	CXMU1	
34:	VCROSSX	
35:	SWEEPK	
36:	CXMU	
37:	RHSY	
38:	VCROSSZ	
39:	CONVRT	
40:	DAMPX	
41:	DAMPY	
42:	SDAMPX	
43:	SPECX	
44:	SDAMPY	
45:	SPECY	
46:	BTRID	
47:	FILLMX	
48:	CHATA	
49:	CHATA	
50:	CHATA	
51:	CHATA	
52:	CHATV1	
53:	CHATV1	
54:	CHATV1	
55:	CHATV1	
56:	FILLHY	
57:	CHATB	
58:	CHATB	
59:	CHATB	
60:	CHATB	
61:	CHATV2	
62:	CHATV2	
63:	CHATV2	
64:	CHATV2	
65:	FILLMZ	
66:	CHATC	
67:	CHATC	
68:	CHATC	
69:	CHATC	
70:	CHATV3	
71:	CHATV3	
72:	CHATV3	
73:	CHATV3	
74:	BTRID	<See 46>
75:	TXIIRHS	
76:	SDIAGX	
77:	SPENTA3	
78:	SPENTA	
79:	SPENTA	
80:	NIRHS	
81:	SDIAGY	
82:	SPENTA3	
83:	SPENTA	
84:	SPENTA	
85:	SWEEPJ2	
86:	CXMU	
87:	SDAMPZ1	
88:	SPECZ	
89:	BTRID	<See 46>
90:	PIRHS	
91:	SDIAGZ	
92:	SPENTA3	
93:	SPENTA	
94:	SPENTA	
95:	TZETRHS	
96:	L2NORM	
97:	SOURCE	
98:	CXMU	
99:	RHSKE	
100:	CXMU	
101:	CXMU	
102:	CXMU	
103:	CXMU	
104:	CXMU	
105:	DAMPKE	
106:	SDAMPZKE	
107:	SPECZ	
108:	SWEEPKE	
109:	SDAMPXYKE	
110:	SPECX	
111:	SPECY	
112:	CXMU	
113:	STEPXYKE	
114:	STEPETAKE	
115:	STEPZETAKE	
116:	CXMU	
117:	SDAMPZ1	<See 87>
118:	L2NORM	
119:	L2NORM	
120:	SST	
121:	INTRBC	<See 9>
122:	BCDRV	
123:	UPDIBC	
124:	GETBND	
125:	NGGRID	
126:	GETQB	
127:	BNDRY	
128:	BNDRYKE	
129:	EVISKE	
130:	CXMU2	
131:	IEDRV	<See 12>
132:	SETPARM	
133:	GST	
134:	OUTPUT	

Call Graph of FDL3DI generated from FORGEExplorer

A FDL3DI Apprentice Results

Routines ordered by inclusive time			Routines ordered by exclusive time		
CHD3D	2725.4200	(97.76)	_SQRT	386.9030	(13.88)
IEDRV	2659.3700	(95.39)	RHSXZ	192.2990	(6.90)
NS3D	2527.5800	(90.66)	SPENTA	169.6670	(6.09)
SWEEPK	1134.6800	(40.70)	VCROX	146.9830	(5.27)
SWEEPJ2	526.6700	(18.89)	SDAMPY	125.8120	(4.51)
SWEEPJ1	503.4750	(18.06)	SDAMPZ	122.3920	(4.39)
_SQRT	386.9030	(13.88)	SPENTA3	117.7680	(4.22)
RHSXZ	283.3470	(10.16)	RHSY	115.6430	(4.15)
SDIAGZ	225.2250	(8.08)	SDIAGZ	103.1130	(3.70)
RHSY	205.9900	(7.39)	SDIAGY	92.5319	(3.32)
SWEEPI	192.8950	(6.92)	VCROSY	90.9966	(3.26)
SDIAGY	188.2610	(6.75)	VCROSZ	90.2956	(3.24)
SDAMPY	174.0400	(6.24)	TXIIRHS	85.8608	(3.08)
SDAMPZ	170.4320	(6.11)	METRIC	84.1194	(3.02)
SPENTA	169.6670	(6.09)	SDAMPX	70.4913	(2.53)
SDIAGX	163.6410	(5.87)	SDIAGX	67.4766	(2.42)
TXIIRHS	159.7770	(5.73)	GST	60.8844	(2.18)
VCROX	146.9830	(5.27)	TZETRHS	60.4157	(2.17)
CXHU	124.9370	(4.48)	SST	59.5777	(2.14)
SDAMPX	117.8230	(4.23)	CHD3D	59.4443	(2.13)
SPENTA3	117.7680	(4.22)	THSTEP	53.5041	(1.92)
SPECZ	96.0547	(3.45)	CXHU	53.1956	(1.91)
VCROSY	90.9966	(3.26)	SPECZ	53.1937	(1.91)
VCROSZ	90.2956	(3.24)	PIRHS	38.8029	(1.39)
TZETRHS	86.0516	(3.09)	SWEEPJ2	33.6610	(1.21)
METRIC	84.1194	(3.02)	NIRHS	29.4131	(1.06)
PIRHS	78.9000	(2.83)	SPECY	26.7436	(0.96)
THSTEP	74.5201	(2.67)	SPECX	25.9076	(0.93)
SDAMPZ1	63.1253	(2.26)	CXHU1	24.0548	(0.86)
GST	60.8844	(2.18)	CONVRT	19.6552	(0.71)
SST	59.5777	(2.14)	SWEEPK	19.0728	(0.68)
NIRHS	50.0930	(1.80)	SDAMPZ1	15.0475	(0.54)
SPECY	48.1981	(1.73)	L2NORM	11.2173	(0.40)
SPECX	47.2684	(1.70)	OUTPUT	5.2283	(0.19)
CXHU1	45.8604	(1.64)	INTRBC	3.6256	(0.13)
CONVRT	19.6552	(0.71)	INITIA	2.7266	(0.10)
L2NORM	11.2182	(0.40)	GETIBK	1.6783	(0.06)
OUTPUT	5.2283	(0.19)	UPDIBC	1.3244	(0.05)
INTRBC	4.6560	(0.17)	GETQB	0.6324	(0.02)
BCDRV	3.2220	(0.12)	PUTQB	0.6227	(0.02)
INITIA	2.7299	(0.10)	BNDY	0.6109	(0.02)
UPDIBC	2.6090	(0.06)	GETINT	0.4061	(0.01)
GETIBK	1.6783	(0.06)	IEDRV	0.2827	(0.01)
GETQB	0.6324	(0.02)	NGGRID	0.2600	(0.01)
PUTQB	0.6227	(0.02)	GETBND	0.2508	(0.01)
BNDY	0.6109	(0.02)	SWEEPJ1	0.0248	(0.00)
GETINT	0.4061	(0.01)	PUTIBK	0.0167	(0.00)
NGGRID	0.2600	(0.01)	PUTINT	0.0039	(0.00)
GETBND	0.2508	(0.01)	PUTBND	0.0025	(0.00)
PUTIBK	0.0167	(0.00)	f\$init	0.0018	(0.00)
PUTINT	0.0039	(0.00)	_ATAN	0.0014	(0.00)
PUTBND	0.0025	(0.00)	SWEEPI	0.0014	(0.00)
f\$init	0.0018	(0.00)	_COS	0.0009	(0.00)
_ATAN	0.0014	(0.00)	_SIN	0.0009	(0.00)
_COS	0.0009	(0.00)	NS3D	0.0009	(0.00)
_SIN	0.0009	(0.00)	SETPARM	0.0006	(0.00)
SETPARM	0.0006	(0.00)	BCDRV	0.0002	(0.00)

(Note: inclusive time includes the time of any called subroutines, exclusive time is the time for a subroutined minus the time spent in any called subroutines or functions).

References

- [1] Juan J. Alonso, Andrey Belov, Scott G. Sheffer, Luigi Martinelli, and Antony Jameson. Efficient simulation of three-dimensional unsteady flows on distributed memory parallel computers. In *Computational Aerosciences Workshop, NASA Ames Research Center*, August 1996.
- [2] Timothy J. Barth. Parallel CFD algorithms on unstructured meshes. In *AGARD Special Course on Parallel Computing in CFD*, May 1995.
- [3] Sukumar Chakravarthy, Oshin Perroomian, and Balu Sekar. A new unified-grid CFD capability for CFD applications on parallel computers. In *Computational Aerosciences Workshop, NASA Ames Research Center*, August 1996.
- [4] Rob F. Van der Wijngaart and Maurice Yarrow. RANS-MP: A portable parallel navier-stokes solver. In *Computational Aerosciences Workshop, NASA Ames Research Center*, August 1996.
- [5] Dennis C. Jespersen and Creon Levit. A data-parallel multizone navier-stokes code. In *Computational Aerosciences Workshop, NASA Ames Research Center*, March 1995.
- [6] Ramki R. Krishnan. An efficient parallelization procedure for multi-block CFD codes. In *Computational Aerosciences Workshop, NASA Ames Research Center*, August 1996.
- [7] Mark J. Lutton and Miguel R. Visbal. Time-accurate validation of the Chimera method for unsteady vortical flows. *AIAA Paper 96-2077*, June 1996.
- [8] Robert L. Meakin. On the spatial and temporal accuracy of overset grid methods for moving body problems. *AIAA Paper 94-1925*, June 1994.
- [9] Kevin Roe and Piyush Mehrotra. Evaluation of High Performance Fortran for CAS applications. In *Computational Aerosciences Workshop, NASA Ames Research Center*, August 1996.
- [10] J.S. Ryan and S.K. Weeratunga. Parallel computation of 3-D navier-stokes flowfields for supersonic vehicles. *AIAA Paper 93-0064*, January 1993.
- [11] Merritt H. Smith and Rob F. Van der Wijngaart. Granularity and the parallel efficiency of flow solution on distributed computer systems. *AIAA Paper 94-2266*, June 1994.
- [12] Merritt H. Smith, Rob F. Van der Wijngaart, and Maurice Yarrow. Improved multi-partition method for line-based iteration schemes. In *Computational Aerosciences Workshop, NASA Ames Research Center*, March 1995.
- [13] N. Suhs and R. Tramel. *PEGSUS 4.0 User's Manual*. Arnold Engineering Development Center, Arnold AFB, TN, June 1991. AEDC-TR-91-8.
- [14] Joseph Vadyak, George D. Shrewsbury, Jim C. Narromore, and Gary Montry. Navier-stokes aerodynamic simulation of the V-22 Osprey on the Intel Paragon MPP. In *Computational Aerosciences Workshop, NASA Ames Research Center*, March 1995.
- [15] Michael Wolfe. *High Performance Compilers for Parallel Computing*. Addison-Wesley, 1996.
- [16] Jerry C. Yan, Dennis Jespersen, and Pieter Buning. Performance evaluation of OVERFLOW/PVM on a network of workstations. In *Computational Aerosciences Workshop, NASA Ames Research Center*, August 1996.

**QUALITATIVE PROCESS MODELING OF CELL-CYTOKINE
INTERACTIONS, ADHESION MOLECULES, AND GENE REGULATION
IN IMMUNITY, INFECTION AND WOUND HEALING**

Robert B. Trelease, Ph.D.
Adjunct Assistant Professor
Department of Neurobiology

UCLA School of Medicine
10833 Le Conte Avenue
Los Angeles, CA 90024-1763

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory,
74th Medical Group, Wright Patterson Air Force Base

September 1996

QUALITATIVE PROCESS MODELING OF CELL-CYTOKINE
INTERACTIONS, ADHESION MOLECULES, AND GENE REGULATION
IN IMMUNITY, INFECTION AND WOUND HEALING

Robert B. Trelease, Ph.D.
Adjunct Assistant Professor
Department of Neurobiology
UCLA School of Medicine

Abstract

The theories and applications of artificial intelligence and qualitative process modeling methodologies were studied in the contexts of cell interactions, cytokines, cell adhesion molecules, and gene regulation processes involved in innate immunity and hyperbaric oxygenation. The principal objective was to develop modeling systems capable of representing basic research experiments, medical treatment processes and outcomes for the purposes of evaluating real experimental data and recursively generating new process rules to extend the existing knowledge base heuristics. System heuristics were successfully created for representing cell interactions, cytokines, adhesion molecules, and gene regulation in immunity, infection and wound healing, and the operations of fundamental qualitative process models were demonstrated in a number of successful simulation experiments.

QUALITATIVE PROCESS MODELING OF CELL-CYTOKINE
INTERACTIONS, ADHESION MOLECULES, AND GENE REGULATION
IN IMMUNITY, INFECTION AND WOUND HEALING

Robert B. Trelease, Ph.D.

"Modeling cannot achieve the rigor of which mathematics is capable. Nor can modeling be done to any predetermined level of precision. But what modeling can do is reduce the need for intuitional insight to manageable bite sizes. The "leaps of faith," so to speak, are kept to jumping over puddles, not across vast canyons. Constantly referring to observations keeps the whole enterprise from drifting into an ethereal world of self-deception. Modeling is an art and always will contain the subjective and cultural values. The best modeling will be done by recognizing these attributes of the trade rather than pretending there is rigor where it cannot be had."

Foster Morrison, The Art of Modeling Dynamic Systems, p. 304

Introduction

Contemporary analyses and models of complex physical systems conventionally rely on quantitative methods (e.g., multiple transfer functions, differential equations, and parametric statistics) to characterize functional relationships between different natural and artificial processes. However, certain classes of model problems deal with incompletely understood processes and/or unquantifiable relationships. To model these problems, qualitative reasoning methods have been developed using cognitive theory and symbolic computer programming techniques employed by artificial intelligence (AI) research modeling of cognitive processes in expert problem-solving. (Weld and De Kleer, 1992).

Knowledge-based qualitative reasoning approaches such as qualitative process (QP) theory (Forbus, 1984) recognize that human experts typically comprehend processes and scientific concepts in qualitative terms (e.g., current increases as voltage rises given a constant electrical resistance) despite the fact that such relationships might be most deterministically described in formal quantitative terms (e.g., $I=E/R$). As used by scientific experts and other problem-solvers, qualitative reasoning can thus be thought of as underlying or preceding the formulation of quantitative descriptions of systems and dynamics (Weld and De Kleer, 1992).

Computer-based qualitative reasoning methods have been successfully used to solve physics problems that are intractable when approached with quantitative techniques (De Kleer and Weld, 1992). Qualitative simulation and modeling have been used for testing system theories, for controlling processes, for diagnosing problems, and for developing "discovery systems" that evolve new process theories based on model behaviors (Weld and De Kleer, 1992). In the realm of process control, QP systems have been used to model and regulate the production of polymer-

based composite materials cured in the complex pressure/temperature/time environment of an autoclave (LeClair and Abrams, 1989; LeClair et al., 1989).

The purpose of the research project described here was to study and to develop basic knowledge-based models of immunological processes involving regulatory molecules (cytokines), cell adhesion molecules, and gene regulation. **The principal focus was the application of QP system modeling methods to support discovery in microbiology and immunology (research on responses to infections) and in hyperbaric oxygenation (HBO), a medical treatment method (in regular use by the Air Force Medical Corps) that uses high-pressure oxygen administration to combat intractable infections and promote accelerated healing of problem wounds.** HBO has a variety of effects on immune system and wound repair processes, including incompletely characterized interactions with cellular inflammatory mechanisms, cytokines, adhesion molecules expression and on regulation of gene expression. Successful implementation of these modeling techniques would make significant contributions to the organization and planning of basic research, to the understanding of complex process interactions in immunity and infection, and to the understanding of healing processes and methods for improving HBO therapy.

Work on hyperbaric oxygenation was performed in collaboration with Colonel Richard Henderson USAF MC, of the Hyperbaric Medicine Unit, 74th Medical Group, Wright-Patterson AFB, Dayton Ohio. Dr. Henderson provided key insights into critical issues and problems in hyperbaric medicine and processes of wound healing, cell adhesion molecules, and gene regulation. Work on cellular immune processes, viral infections, cytokines and cell adhesion molecules was performed in collaboration with Dr. Nancy Bigley, Department of Microbiology and Immunology, Wright State University School of Medicine. Assistance with the knowledge base construction, rule writing, and general debugging for the cellular immune processes, viral infections, cytokines and cell adhesion molecules projects was provided by Jennifer A. Raker, a graduate student in Microbiology and Immunology at Wright State University School of Medicine. Ms. Raker was a participant in the AFOSR Summer Research Program and provided outstanding support of these projects, in addition to performing brilliantly in her own research work. Assistance with qualitative process and discovery theories and system internals was obtained from Jack Park, developer of the symbolic programming tools used for this work. The projects described here are significant expansions on previous AFOSR SFRP/SREP work described in Trelease (1994) and Trelease and Park (1996).

Methodology and Discussion of Problem Domain Issues

All knowledge based development and programming work for this project was performed on Apple Macintosh computer systems with 10 or more megabytes of random access memory and 105 megabyte or larger hard disks. All software systems development was performed in a Macintosh-based symbolic programming environment constituting an extensible, knowledge-based, qualitative process modeling system. System development was rooted

in accepted practices and theories in AI, cognitive sciences, epistemology, and scientific discovery. In Section I, below, we discuss some of the basic tenets and methodology of AI discovery systems. In Section II, we outline fundamental considerations for building discovery systems using The Scholar's Companion. In Section III, we examine specific problem domain issues for immune system cellular interactions, cytokines, cell adhesion molecules and for the gene regulatory effects of reactive oxygen species (ROS) that may be evoked by hyperbaric oxygen treatment.

I. Methods for Knowledge-Based Systems and Qualitative Process Modeling

Kocabas (1991) has distinguished four loosely defined, overlapping types of knowledge necessary for research and discovery: Commonsense, technical, theoretical, and methodological. All of these types are currently computationally representable with symbolic programming methods.

In expert systems and related knowledge-based systems, knowledge is typically represented in the form of heuristics or production rules composed in an object-oriented symbolic programming language (typically a LISP dialect). The collection of rules for a given system is referred to as a knowledge base (KB), and it encompasses theories, hypotheses, and functional descriptions for entities and processes for the problem/model domain. Inferences about real-world problems can be made by an expert system's inference engine, which evaluates data (i.e., from external processes) via forward- or backward-chaining application/evaluation of production rules.

Early work with expert systems focused on strategies for supporting diagnosis (identification of faults and problems) via the interaction of simulation and inference (Davis, 1983). One such method was termed discrepancy detection and relied on handling discrepancies between simulation inferences and real device data. Given the context of diagnosis, discrepancy detection thus qualified as a strategy for discovery.

QP reasoning methodology, in fact, emerged from expert systems research directed at emulating related problem-solving. While expert systems demonstrated domain-specific and global limitations related to the need for extremely large KBs to represent adequately real experts' knowledge, useful qualitative modeling tasks and problem-solving could be performed with relatively compact KBs. In the context of discovery, QP models might be effectively employed in systems for evaluating experimental data and evolving new model heuristics (hypotheses).

In order to function effectively in process discovery, QP systems must possess some special characteristics related to handling functional actor and state changes. Discrete states must be recognized as changes occur. Models must be general enough to allow for variations in real world data, but expectation failures must be recognized and handled in order to generate discoveries in the form of new process heuristics. Automated (heuristic-generating) discovery systems can work to produce new theories by induction, deduction, or abduction.

Practical discovery systems remain somewhat like the "Holy Grail" of some AI researchers, and relatively few working systems exist (Kocabas, 1991). The effort described in this report was directed primarily at generating a robust model that might serve as the core of an experimental discovery system, although all discovery processes involved human intervention in the form of recognizing disparities between TSC system and real world behaviors and then creating new process rule heuristics.

II. Development System Issues and The Scholar's Companion

Development for this project was undertaken using The Scholar's Companion (TSC; ThinkAlong Software, Brownsville, CA) a symbolic programming environment composed of nested, extensible multiple language interpreters/compiler, run-time utilities, an envisionment builder, and basic tools supporting expectation definitions, failure handling, data file and database loading, device/process control, envisionment study, and guided process rule mutation.

A. Knowledge Bases, Tasks, and Envisionments

Comparable to a more conventional AI expert system, TSC allows the definition of actors (e.g., different types of cells), taxonomic and functional relationships, physical and biochemical processes, and predicated interactions and behaviors in hierarchical sets of symbolic production rules. TSC also compiles the description of initial conditions and functional designs for simulated experimental trials to be run using defined KB actors (cells) and process behaviors. As distinguished from an expert system that uses a logical inference engine to evaluate acquired data in the context of logical relationships defined by the KB, TSC employs an "envisionment builder" to record and to visualize the evolution of process behaviors and cell interactions resulting from initial experimental conditions given the defined actors, substances, processes, and functional relationships. An envisionment is thus the depiction of all possible behaviors that can be produced given the initial conditions of a simulation/experiment and the specific actors, states, relations, predicates, and process rules defined in the related KB. Envisionments are represented by the output of "log files" that record all the antecedents, consequent terms, and super-episodes for all episodes (discrete system states encountered) as well as by "browser" graphics depicting the order and ontologic relationships between episodes (see Figures 4-10).

TSC performs its functions under the direction of TASK statements. FILLIN tasks, for example, set up the system behaviors necessary to perform a simulation based on actors, states, relations, predicates, and process rules defined in the KB. Other tasks can be created to set up system behaviors to study specific envisionments. STUDY tasks can result in the identification of the sequence of process rule firings necessary for achieving a desirable process outcome. Other study tasks can be created to compare real world data with "expectations" defined in envisionment (process simulation) episodes, a crucial behavior in setting up TSC to produce discovery about processes. Rule mutation

behaviors can be employed following study tasks to produce new process rules that can be used in building new envisionments. Finally, control tasks can be set up to handle process control based on process simulation envisionments.

D. Explanation, Rule-building Strategies, and Building New Envisionments

Explanation seeks to account for system behaviors in terms of defined process heuristics. When new system behaviors are encountered, explanation must involve building of new processes rules to account for those unexpected behaviors. This involves rule-building heuristics that build on or "mutate" existing process rules, in addition to designing system tasks that study experimental data and generate new envisionments based on new rules.

III. Problem Domain Issues

The work described in this report builds on immune system knowledge bases first developed for modeling cell-cell-cytokine interactions during infectious processes (AFOSR SFRP/SREP, 1994-1995). That specific formal immune system model was composed of three conceptual parts: 1) A core biological knowledge base (KB) describing actors, taxonomic relationships, process definitions and predicates, 2) sets of behavioral rules for immune system actors, pathogens, and processes, and 3) sets of rules and conditions for the execution of experimental simulations. All rules and concepts (actors, predicates, etc.) were defined as frames in a Scheme dialect of the LISP programming language, and model components were divided into a group of loadable text files. This modular approach was taken in order to support orderly expansion of the KB and problem domain with additional heuristics, for example, for molecular biology, additional histology, pathological processes and treatments.

This core biological KB was composed of three incrementally compiled files. BIOPRIMITIVE.T included definitions of underlying physical, chemical, and biological actors (e.g., atoms, molecules, proteins, cells, organisms, endocytosis), taxonomic definitions and inheritance (e.g., class, order, family, genus, species), and some basic predicate relationships between various actors. CELL.T included fundamental definitions of biological structures (e.g., nuclei, membranes, mitochondria, etc.) substances and predicates central to cell biology. IMMUNE.T was comprised of definitions of immune system cells, receptors, cytokines and other actors, substances, organs, and predicates.

These specific heuristics for immune system cell processes, interactions, and behaviors were contained in IMMUNE.RBT. Different sets of initial conditions (specific actors, predicates and relationships) and specific experimental designs were maintained in separate files (e.g. IMMUNE.EXP1, IMMUNE.EXP2, etc.) allowing different experimental simulations to be run with the same underlying biological/immunological KB. All concepts and process rules were derived from current theories and objective information published in the current general

biological (Purves et al., 1992), systematic (Fortuner, 1993), histological and cell biological (Junquiera et al., 1992; Weiss, 1993; Goodman, 1994) and specialized immunological (Abbas et al., 1991; Janeway and Travers, 1996) literatures.

Consistent with this modular approach and the principle of incremental KB expansion, the work for the current (summer 1996) project involved adding many new actors and process heuristics to the aforementioned files. Numerous new actors and predicates were added to BIOPRIMITIVE.T, CELL.T, and IMMUNE.T to support an expanded taxonomy and ontology of regulatory molecules, states, relationships and predicates for cytokines, cell adhesion molecules, enzyme regulation, and multiple gene transcription/regulation processes. Three different "subsets" of cellular/immune processes were modeled in order to yield discrete simulations of different functionalities: Cytokine interactions in differential gender-related innate immune responses to viral infection, cell adhesion molecule expression changes in innate immune responses, and gene regulation (protooncogene transcriptional control by antioxidants, cytokines and other bioactive molecules).

A. Cytokine Interactions in Differential Gender-related Innate Immune Responses to Viral Infection

In a previous AFOSR Summer Faculty Research project, we began creating a model of cytokine secretions during innate immune responses to encephalomyocarditis virus (EMCV) infections in male and female mice. This work was performed in collaboration with Dr. Nancy Bigley, Department of Microbiology and Immunology, Wright State University School of Medicine, and was based on the findings of laboratory experiments performed by her graduate student Rafael Curiel (Curiel, 1994; Curiel et al., 1993). Essentially, different pathologies were induced by different types (genetic variants) of EMCV (Cerutis et al., 1989), with females tending to be more resistant than males to organ damage and death evoked by specific virus types. Secretion of specific cytokines at during given time intervals was shown to be significantly different in males and females, and this was assumed to be a primary factor in gender-related morbidity and mortality. To achieve initial results with the QP model, it had been necessary to define elapsed time as a specific actor in process rule heuristics.

For the current project, the regulatory substances taxonomy was greatly expanded in order to include steroids, neuromodulators, and other substances that might be effectors in processes that could account for gender-related susceptibility to specific pathophysiologic effects; part of this taxonomy is depicted in Figure 1. We refined the use of qualitative time actors in the model and included new process rules for inhibitory cytokine interactions. To achieve the most efficient envisionment, we implemented "qualitative intervalometer" heuristics that would allow firing of all possible rules relevant to a given time period before advancing to the next time interval.

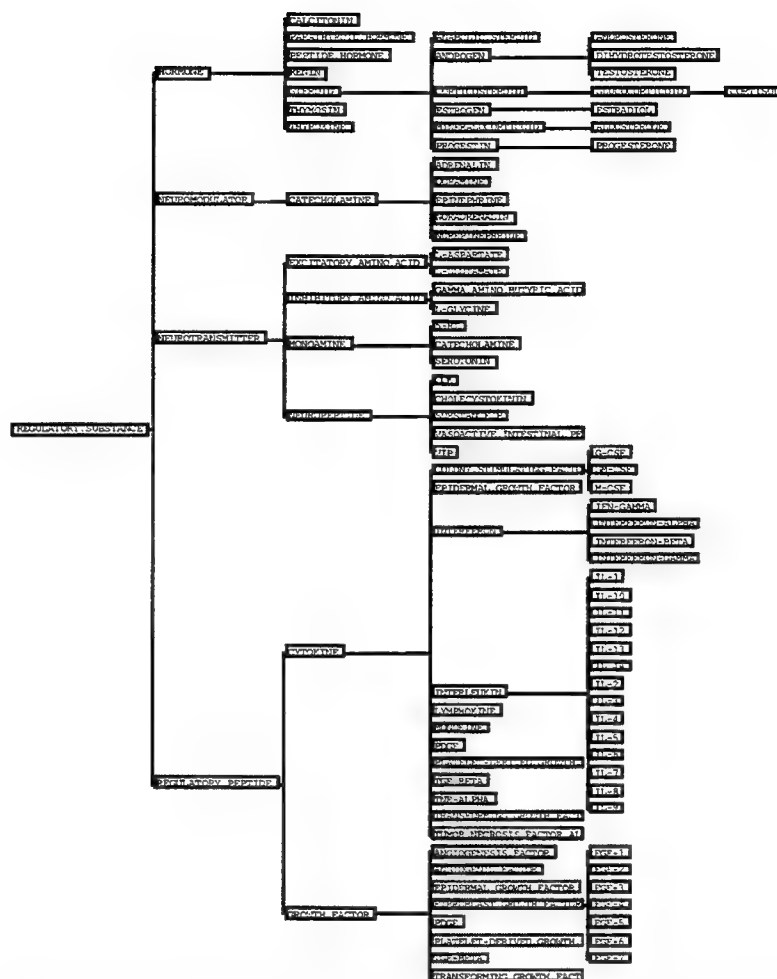


Figure 1: Graphic depiction of taxonomic relationships between some of the cell regulatory substances in the knowledge base

B. Cell Adhesion Molecule Expression Changes in Innate Immune Responses

During immune responses, tissue development and regeneration, special proteins called cell adhesion molecules (CAMs), expressed on cell surfaces, mediate signaling, binding, regulation of cytoplasmic and nuclear and transport processes. Different combinations of CAMs bind and regulate contacts between different and similar cell types. In immune system function, CAMs are responsible for lymphocyte migration behaviors, cell activation, and tissue specific responses (Imhof and Dunon, 1995; Janeway and Travers, 1996). During development, the differential expression of certain CAMs may control the development of specific tissues, structures, and organs. Tissue reconstruction during healing may involve the expression of molecules that stimulate the production of ground substance proteins (e.g., collagen) and promote the growth of connective tissue.

CAMs ontology and taxonomy were created by adding new actors, states, relations, and predicates to the pre-existing KB files as described above. Figure 2 depicts taxonomic relationships between these new actors. The major CAM

families, immunoglobulins, integrins, and selectins, were represented along with their components and structurally based subfamilies.

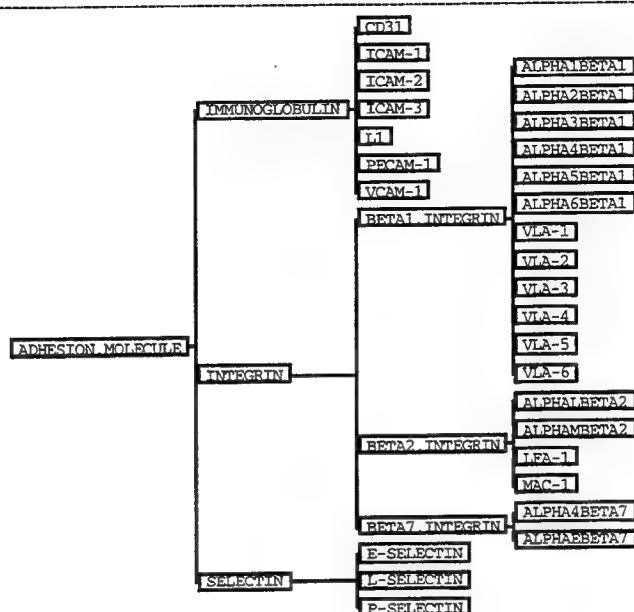


Figure 2: Graphic depiction of taxonomic relationships between some of the cell adhesion molecules in the knowledge base.

Process rule sets were created for representing differential CAMs expression on a variety of immune system, vascular, and organ cell types according to specific cell/system states, based on current knowledge as described in Janeway and Travers (1966). Several experiments were designed to demonstrate differential expression of different CAMs on immune system cells and on endothelia (cells lining the blood vessels) of several organs affected by different strains of the EMCV in primary immune responses, in a different view of correlates of the EMCV infections discussed above. Attention was paid to representing the circulation of "naive" T-lymphocytes, their entry into lymph nodes, subsequent activation by antigen presenting cells, then return to the general circulation where they could pass to affected organs. This "generic" model was intended to serve as the basis for an expanded modeling system capable of representing the different histopathologies known to accompany infection by specific viral strains in both males and females, for purpose of providing some outcome predictions for actual laboratory (in-vivo) experiments.

C. Gene Regulation: Protooncogene Transcriptional Control, Antioxidants, and Cytokines

Certain classes of evolutionarily conserved genes are transcribed when a wide range of cellular processes occur, ranging from activation and cytokine release in immune system cells, through protein-synthesis and signal transduction involved in tissue reconstruction/healing. By evolutionarily conserved we mean genes that control fundamental cellular processes across a wide range of phylogenetic Phylae and Classes (taxa) with very little change

in structural (DNA) sequences. These regulatory sequences, originally identified as "protooncogenes" (due to their association with growth regulation in cancerous cells) are thus ubiquitous in biological systems and thus very important to the understanding of the control of many crucial life processes.

Much recent evidence indicates that two classes of such genes, regulated by NF- κ B (nuclear factor kappa B) and activation protein 1 (AP-1) proteins may be transcriptionally regulated by reactive oxygen species (ROS) at various steps in cellular proteolytic and signal transduction pathways. These gene regulatory processes may thus be particularly important to modeling and understanding immunologic, other antimicrobial, reconstructive and other healing mechanisms enhanced by high pressure oxygen administration in HBO. Shortly after the beginning of this summer research program, Sen and Packer (1996) published a comprehensive paper describing an integrative conceptual model accounting for the complex effects of ROS at multiple steps in the AP-1 and NF- κ B regulatory pathways within cells and nuclei. Figure 3 graphically depicts many of the essential elements of their conceptual model.

The principal objective of this part of the project was to implement a computer-based qualitative process model of this basic Sen and Packer conceptual model. The regulatory substances ontology and taxonomy were greatly expanded by adding new actors, states, relations, and predicates to the pre-existing KB files as described above. Furthermore, process rule sets were created for representing ROS-related interactions with cell enzymes and other regulatory process steps involved in AP-1 and NF- κ B mediated transcription of several generic promoter/enhancer genes. Several experiments were designed to demonstrate different outcomes of these gene regulatory processes in the presence and absence of different factors (ROS and AP-1 and NF- κ B regulatory proteins). This "generic" model was intended to serve as the basis for an expanded modeling system capable of representing changes in gene regulatory processes associated with the production of specific cytokines, cell behaviors, and other outcomes characteristic of HBO and other tissue processes.

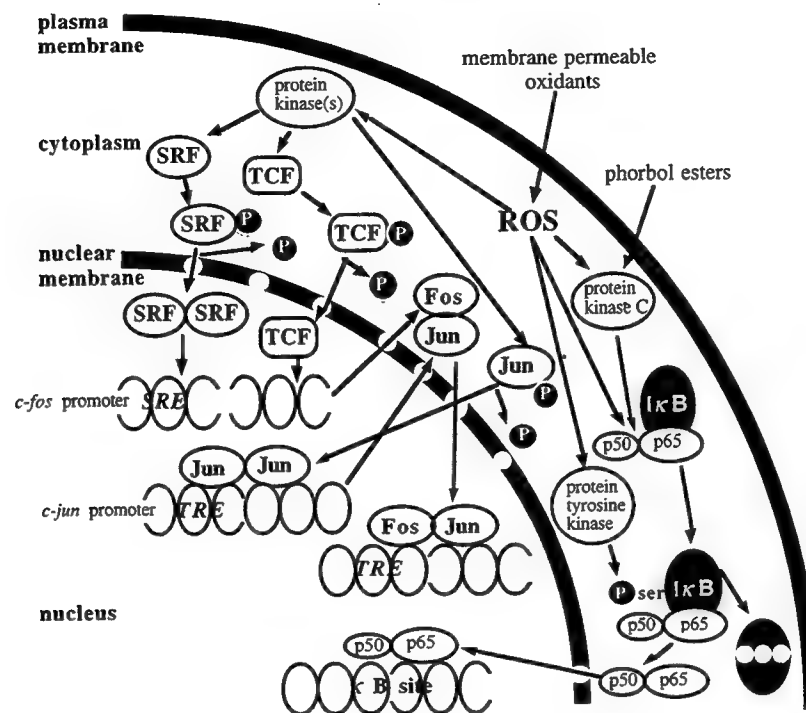
Results

A. Cytokine Interactions in Differential Gender-related Innate Immune Responses to Viral Infection

Initial visualizations generated without "qualitative intervalometer" heuristics were very massive (greater than 3500 unique episodes) and took over three weeks apiece to build on a very fast, RISC-based Macintosh (132 MHz, 64 bit PowerPC 604 processor using 256 Mb RAM). Almost invariably, they exceeded the available application parameter memory and were thus incomplete when they halted. This was apparently due to excessive branching of process paths associated with process rules based on independent time actors for each specific cell and cytokine actor. Because of this apparent limitation, the use of time in the model was reconsidered, with the resulting development of the qualitative intervalometer process and refinement of cytokine secretion heuristics for characteristic immune system cells.

Figure 3:

Graphic depiction of the hypothetical scheme for steps in cellular NF- κ B and AP-1 regulatory genes (protooncogene) activation that may be influenced by oxidants (ROS-reactive oxygen species) and antioxidants (after Sen and Packer, 1996)



These revised heuristics resulted in very tractable environments encompassing several hundred episodes apiece, and clearly demonstrating gender- and viral-strain-related differences in cytokine secretion processes according to time interval. Environments could be built in a few hours on the PowerMac 9500, and the demonstrated cytokine secretion process differences qualitatively resembled those reported for in-vivo experiments published in the scientific literature.

Different environments were produced for male and female subjects, with infection with either MM or D variants of the EMCV virus: Four different environments were produced, as shown in Figures 4 through 7. Male models produced interleukin 10 (IL-10) but either a low concentration (D variant) of or no (MM variant) gamma interferon (IFN- γ) at 12 hours post-infection. Female models produce a high concentration of IFN- γ at 12 hours.

processes hypothesized to occur in EMCV variant infections known to result in myocardial lymphocytosis and histopathology (Cerutis et al., 1989).

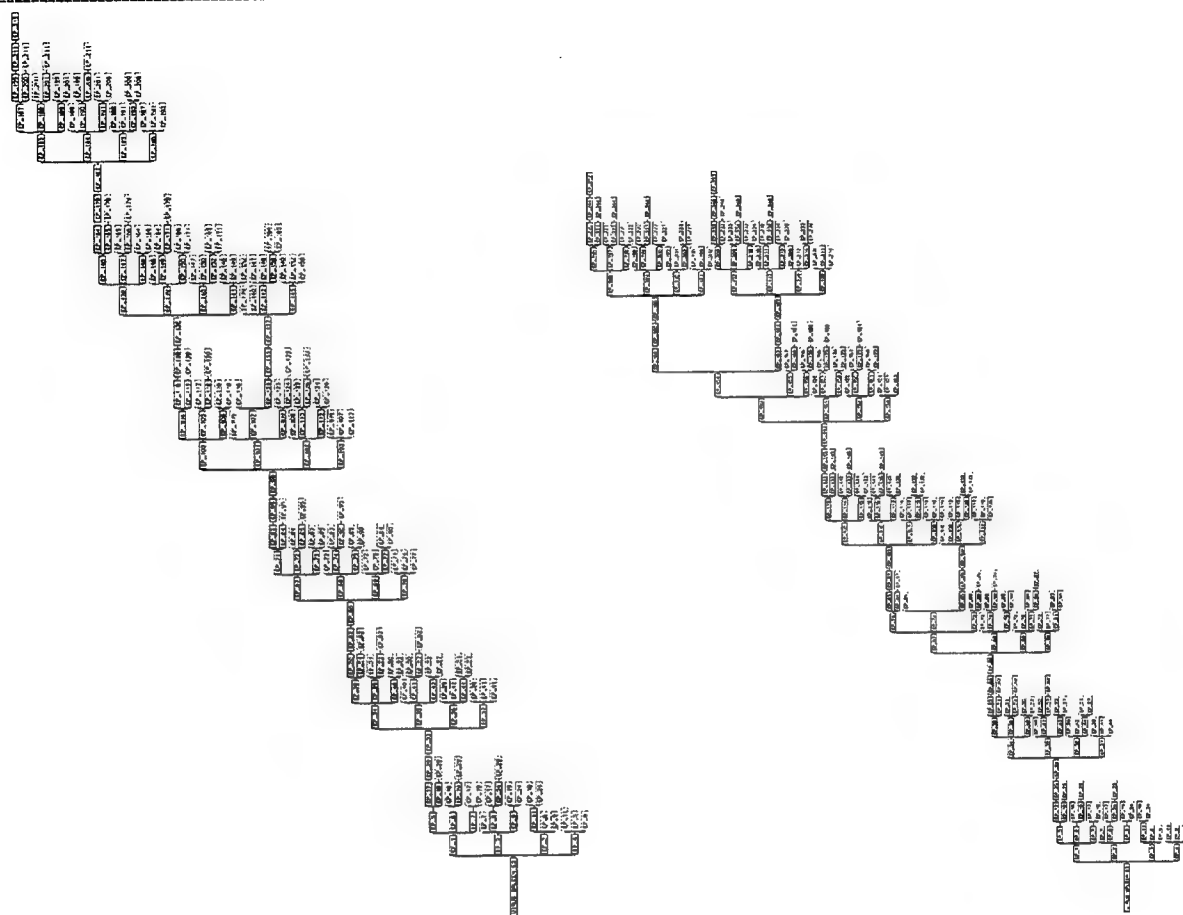


Figure 5: Left--Browser graphic depiction of an environment demonstrating different cytokines secreted at 3, 6, 9, 12, and 24 hours during an innate immune response following EMCV (MM variant) in a male mouse. (Figure rotated 90 degrees CCW due to space limitation)

Figure 6: Right- Browser graphic depiction of an environment demonstrating different cytokines secreted at 3, 6, 9, 12, and 24 hours during an innate immune response following EMCV (D variant) in a female mouse. (Figure rotated 90 degrees CCW due to space limitation)

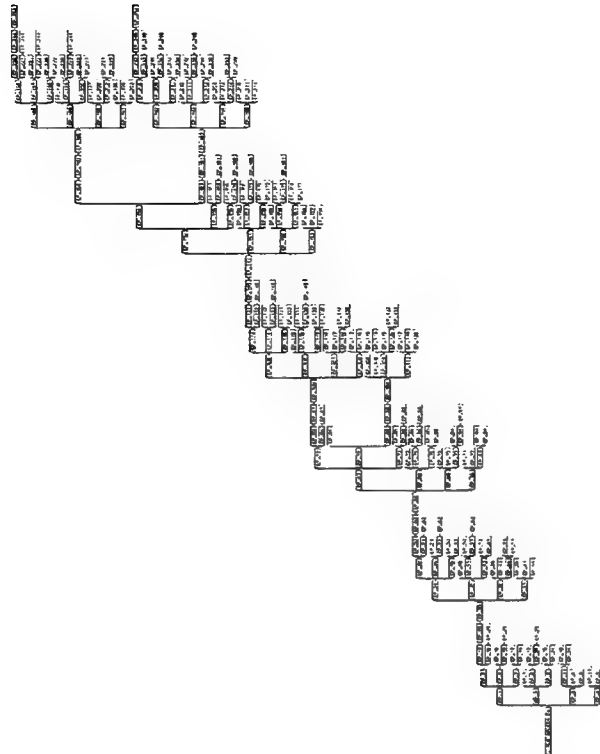


Figure 7: Browser graphic depiction of an envisionment demonstrating different cytokines secreted at 3, 6, 9, 12, and 24 hours during an innate immune response following EMCV (MM variant) in a female mouse. (Figure rotated 90 degrees CCW due to space limitation)

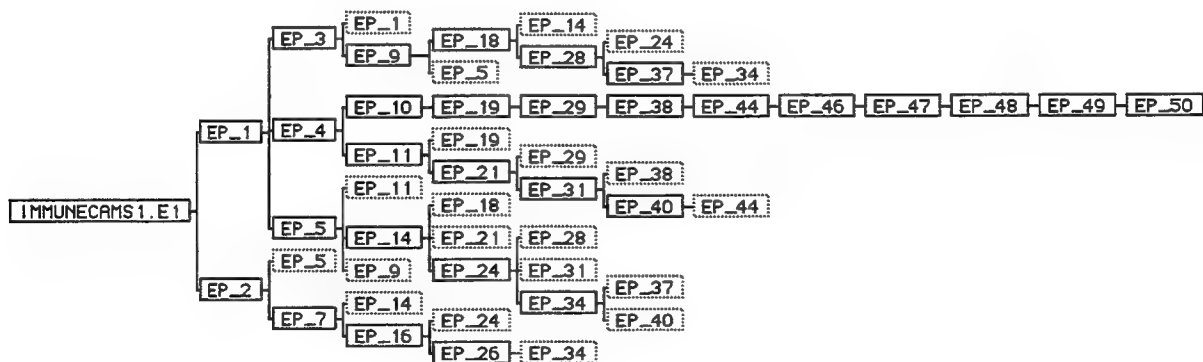


Figure 8: Browser graphic depiction of an envisionment demonstrating changes in CAMs expression, molecule-specific cell binding processes, associated with general circulation, diapedesis in to a lymph node, t-lymphocyte activation, return to the general circulation , and target-organ migration in innate immune responses to EMCV.

C. Gene Regulation: Protooncogene Transcriptional Control, Antioxidants, and Cytokines

The basic gene regulation model experiment environments demonstrated the fundamental transcriptional promoter processes known to occur with oxidant activation of phosphorylation (protein kinase) enzymes (according to Sen and Packer, 1996). Depending on initial conditions, such as the absence of a promoter promoter or the a full complement of necessary agents, the model would produce environments showing partial activation (Figure 10) or activation of a full complement of the appropriate genes (Figure 9). The fundamental conceptual model of Sen and Packer was full implemented at their chosen level of abstraction for the generic effects of ROS on intracellular enzymes catalyzing the activation of regulatory proteins.

Conclusions and Discussion

During the period of this Summer Research Program, fundamental simulations were successfully run for each of the three different "subsets" of cell regulatory processes representations, demonstrating the practicality of qualitative modeling methodology for characterizing functional interactions for different aspects of cell biology. Although these simple discrete simulations viewed coexisting processes separately within different frames of reference, our QP modeling environment would permit the development of more complex, multi-modal simulations involving multiple, nested tasks that could represent processes occurring at superficial and deep levels of abstraction for purposes of explanation of gross behavioral as well as subcellular aspects of immune responses and wound healing. There is certainly plenty of room for expanding these basic simulations and "fleshing out" the modeling system so it more faithfully and comprehensively reflects the functions and interactions of known biological processes.

There are several avenues of continued development for these modeling systems that will be pursued in the coming years. The EMCV/cytokine-interactions model will continue to be expanded as new laboratory data refine our knowledge of what specific processes actually occur, with factual heuristics replacing those that were created by conjecture. Refined model environments will be used to predict potential outcomes of new experiments, with further iterative testing of these conjectures. Work will be done to design a "closed-loop" system for conjecturing outcomes, loading and evaluating actual laboratory experimental data, and automatically building new heuristics from expectation failures and unanticipated outcomes.

In the realm of gene regulation models, fundamental ("generic") heuristics will be expanded to include representation of processes entailed in the published findings of experiments detailing the promotion of AP-1/NF- κ B for specific proteins (e.g., tumor necrosis factor; TNF α) (Kopp and Ghosh, 1995; Sen and Packer, 1996) in response to oxidative stimulation (i.e., such as that encountered in HBO). Candidates for such specific representation include TNF α , IFN- γ , and IL-1.

Because CAMs have widespread involvement in manifold immune and other cell regulatory processes, there are many ways in which the basic modeling system could be expanded. We propose to focus on completing the representation of processes involved in innate immunity to ECMV, with specific heuristics aimed at elucidating mechanisms that may be operational in effecting the organ-specific pathologies known to be caused by different genetic variants of that virus (Cerutis et al., 1989).

References and Bibliography

Abbas, A.K., Lichtman, A.H., and Pober, J.S. Cellular and Molecular Immunology. W.B. Saunders Company: Philadelphia, 1991.

Andersson, E.C., Christensen, J.P., Marker, O., and Thomsen, A.R. Changes in cell adhesion molecule expression on T cell associate with systemic virus infection. Journal of Immunology 152:1237-1245, 1994.

Beiqing, L., Chen, M., and Whisler, R. Sublethal levels of oxidative stress stimulate transcriptional activation of c-jun and suppress IL-2 promoter activation in Jurkat T cells. Journal of Immunology 157:160-169, 1996.

Cerutis, D., Bruner, R., Thomas, D., and Giron, D. Tropism and histopathology of the D, B, K, and MM variants of encephalomyocarditis virus. Journal of Medical Virology 29:63-69, 1989.

Christensen, J.P., Andersson, E.C., Scheynius, A., Marker, O., and Thomsen, A.R. Alpha4 integrin directs virus-activated CD8-T cells to sites of infection. Journal of Immunology 154:5293-5301, 1995.

Curiel, R.E. Cytokines Expressed By Cultured Splenocytes During The First 24 Hours Of Picorna Virus Infection Reflect The Disease Susceptibility Or Resistance Of The Spleen Cell Donors. Unpublished doctoral dissertation, Wright State University, Ohio, 1994.

Curiel, R.E. , Miller, M., Ishikawa, R, Thomas, D. and Bigley, N.J. Does the gender difference in interferon production seen in picornavirus-infected spleen cell cultures from ICR Swiss mice have any in vivo significance? Journal of Interferon Research 13:387-395, 1993.

Davis, J.C. and Hunt, T.K. Problem Wounds. The Role of Oxygen. Elsevier: New York, 1988.

De Kleer, J. and Forbus, K. D. Building Problem Solvers. MIT Press: Cambridge, MA, 1993.

De Kleer, J. and Weld, D.S. Qualitative Physics: A Personal View. In: Weld, D.S. and De Kleer, J. (Editors). Readings in Qualitative Reasoning About Physical Systems. Morgan Kaufmann: San Mateo, 1992. Pp 1-8.

Forbus, K. D. Qualitative process theory. Artificial Intelligence 24:85-168, 1984.

Fortuner, Renaud (Editor). Advances in Computer Methods for Systematic Biology. The Johns Hopkins University Press: Baltimore, 1993.

Goodman, S.R. Medical Cell Biology. J.B. Lippincott Co.: Philadelphia, PA, 1994.

Hemler, M.E. and Mihich, E. (Eds.) Cell Adhesion Molecules: Cellular Recognition Mechanisms. Plenum Press, New York, 1993.

Hyperbaric Oxygen Committee. Hyperbaric Oxygen Therapy: A Committee Report. Undersea and Hyperbaric Medical Society: Bethesda, MD, 1992

Imhof, B.A. and Dunon, D. Leukocyte migration and adhesion. Advances in Immunology (58):345-416, 1995.

Janeway, C.A., and Travers, P. Immunobiology. The Immune System in Health and Disease. Current Biology, Ltd./Garland Publishing, Inc: London, 1996.

Junqueira, L.C., Carneiro, J., And Kelly, R.O. Basic Histology. (7th Edition). Appleton and Lange: Norwalk, CT, 1992.

Karp, P.D. Hypothesis formation as Design. In: Shrager, J. and Langley, P. (Editors). Computational models of scientific discovery and theory formation. Morgan Kaufman: San Mateo, CA, 1990. Pp. 276-317.

Kloth, L.C., McCulloch, J.M., and Feedar, J.A. (Editors). Wound Healing: Alternatives in Management. F.A. Davis Company: Philadelphia, 1990

Kocabas, S. Computational models of scientific discovery. The Knowledge Engineering Review 6(4):259-305, 1991.

Kopp, E.B., and Ghosh, S. NF-kB and Rel proteins in innate immunity. Advances in Immunology 58:1-27, 1995.

LeClair, S.R., Abrams, F.L., and Matejka, R.F. Qualitative process automation: Self-directed manufacture of composite materials. Artificial Intelligence for Engineering Design, Analysis and Manufacturing. 3(2):125-136, 1989.

LeClair, S.R., Abrams, F.L. Qualitative process automation. International Journal of Computer Integrated Manufacturing 2(4):205-211, 1989.

Morrison, F. The Art of Modeling Dynamic Systems. Forecasting for Chaos, Randomness, and Determinism. John Wiley & Sons, Inc.: New York, 1991.

Nordhausen, B. and Langley, P. An integrated approach to empirical discovery. In: Shrager, J. and Langley, P. (Editors). Computational models of scientific discovery and theory formation. Morgan Kaufman: San Mateo, CA, 1990. Pp. 98-128.

Purves, W.K., Orians, G.H., Heller, H.C. Life, the Science of Biology, Third Edition. Sinauer Associates: New York, N.Y., 1992.

Sen, C.K. and Parker, L. Antioxidant and redox regulation of gene transcription FASEB Journal 10:709-720, 1996.

Shrager, J. and Langley, P. Computational approaches to scientific discovery. In: Shrager, J. and Langley, P. (Editors). Computational models of scientific discovery and theory formation. Morgan Kaufman: San Mateo, CA, 1990. Pp.1-25.

Shrager, J. and Langley, P. (Editors). Computational models of scientific discovery and theory formation. Morgan Kaufman: San Mateo, CA, 1990.

Trelease, R.B. Development of a feature-based knowledge base for biologically-based materials and processes. Contributed Research and Development 145:1-10, 1994.

Trelease, R.B. and Park, J. qualitative process modeling of cell-cell-pathogen interactions in the immune system. Computer Methods and Programs in Biomedicine, In Press, 1996.

Weiss, Leon. Histology. Cell and Tissue Biology. Fifth Edition. Elsevier Biomedical: New York, 1983.

Weld, D.S. and De Kleer, J. (Editors). Readings in Qualitative Reasoning About Physical Systems. Morgan Kaufmann: San Mateo, 1992.

**SIMULATION OF ANTI-JAMMING GPS ARRAYS USING
FINITE ELEMENT SOFTWARE**

John L. Volakis¹ and Mark Casciato²

¹Professor ²Graduate Research Assistant

Radiation Laboratory

Dept. of Electrical Engineering and Computer Science

**University of Michigan
1301 Beal Ave.
Ann Arbor, MI 48109-2122**

Final Report for:

Summer Faculty Research Program

Wright Laboratory

Contact: Denice Jacobs and Dana Howell

Code: WL/AAST, Phone: (513) 255-2766

Wright Patterson AFB, OH 45433

Sponsored by:

Air Force Office of Scientific Research

and

Wright Laboratory, WPAFB

August 1996

SIMULATION OF ANTI-JAMMING GPS ARRAYS USING FINITE ELEMENT SOFTWARE

John L. Volakis¹ and Mark Casciato²

¹Professor ²Graduate Research Assistant

Radiation Laboratory

Dept. of Electrical Engineering and Computer Science

University of Michigan

Abstract

Two sophisticated anti-jam antennas explicitly designed for the Global Positioning System (GPS) satellite communications were simulated and analyzed using the University of Michigan finite element software FEMA-PRISM. The software was upgraded to allow simulation of these unique GPS antenna arrays which consisted of a reference element and an array of six anti-jam auxiliary elements that produce pattern nulls at the direction of the incoming jamming signal. The goal of the study was to examine the effectiveness of current state-of-the-art software for simulating complex antenna arrays and to generate accurate volumetric patterns (incorporating mutual coupling, losses, dispersion, etc.) of the actual array to be used in a hardware simulator when evaluating the anti-jam performance of the antenna array under test. Our study demonstrated that the University of Michigan finite element code FEMA-PRISM with (minor modifications) was capable of analyzing the two GPS antennas without compromising their geometrical features in performing the numerical analysis. For one of the arrays, the antenna element was a slot backed by a triangular cavity whereas for the second array the basic element was a dual patch to accommodate radiation at the L1 (1575 MHz) and L2 (1227 MHz) GPS bands. Patterns and input impedance curves were generated for each antenna and these were validated with reference data whenever possible. With the availability of the simulation models and discrete geometry data for each antenna, future efforts will focus on evaluating the performance of the arrays under various jamming conditions and realistic GPS operational scenarios.

SIMULATION OF ANTI-JAMMING GPS ARRAYS USING FINITE ELEMENT SOFTWARE

John L. Volakis and Mark Casciato

Introduction

Over the past 7 years, the University of Michigan has been developing computational methods and related computer codes for the analysis of antennas on complex platforms. Antennas are the "eyes" and "ears" of all communication devices, including radar, microwave imaging systems, intelligent vehicle guidance systems and for wireless communication devices for cellular, personal communication systems (PCS), global positioning system (GPS) and local area networks (LAN) applications. The latter has been a major driver toward the design, development and deployment of suitable antennas which are efficient and satisfy the requirements of personal, military and commercial systems.

In the past, antenna design was based on a few well known geometrical shapes such as patches, horns, reflectors, spirals, log-periodic, Yagi-Uda (used for TV reception), wire dipoles (typically used for car radios) and wire loops. Some of these are known to be broadband (for example, log-periodic) and others are known to have very narrow bandwidth (for example, the patch). Designers usually worked with these classic shapes and chose their geometrical parameters to satisfy the frequency, bandwidth and gain pattern requirements. Deviation from the classical shapes requires substantial investment in experimental and analytical research and development. Increasing and specialized requirements on pattern, bandwidth and gain require new antenna designs as well as feed networks. Indeed, the cost of developing new antenna designs and analysis by experimental means is prohibitive and over the past few years there has been an impetus to develop antenna analysis and design software. At the University of Michigan we developed a unique code referred to as FEMA-PRISM for conformal antenna analysis [1]. This code is the first to be suitable for simulating antennas on doubly curved platforms. Moreover, it is based on a new hybrid finite element methodology which permits the analysis of antennas of varying shape and material composition provided the geometrical data are supplied in a certain form. Additional information about FEMA-PRISM and its formulation can be found in the references [2-11]. A brief description of the University of Michigan finite element analysis codes is given in the Appendix.

Scope of the Work

The scope of our work over the past two months has been the application and further development of FEMA-PRISM to two GPS antennas with anti-jamming auxiliary antenna elements in support of the Wright Laboratory anti-jamming antenna array simulator [12]. This hardware simulator evaluates the anti-jamming performance of various antenna arrays (over a frequency band and as a function of reception angle). A typical antenna element along with anti-jamming array elements is depicted in Figure 1. Presently, the simulator assumes that each of the antenna elements is radiating as isotropic source (i.e., it is assumed to radiate equally in all directions), whereas in practice

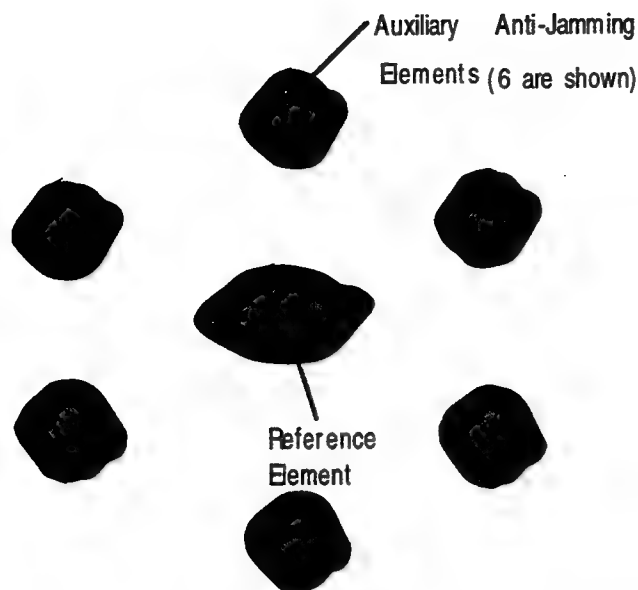


Figure 1. Illustration of the anti-jamming array. The reference element is the main receiver and the excitation of the auxiliary elements are adjusted to generate a null at the direction of the jamming signal.

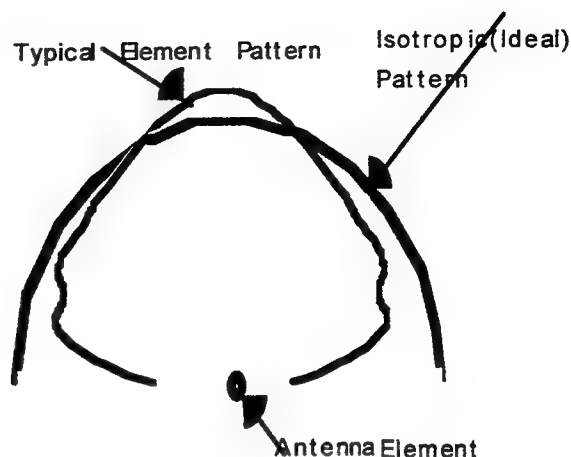


Figure 2. Comparison of radiation patterns due to an isotropic source and that of a typical actual element.

the antenna pattern can deviate substantially from the isotropic pattern conditions as illustrated in Figure 2. The actual pattern depends on the shape and geometry of the element as well as the shape of the platform (aircraft, missile and so on). Furthermore, the anti-jamming element patterns and their coupling with the reference element (or among themselves) present additional variables which must be considered in the performance of the array in the presence of jamming signals. Typically, the anti-jamming elements are used to generate a canceling signal (nulling) at specific directions (the jamming signal directions) using adaptive signal processing algorithms. This cancellation must be accomplished with extreme accuracy and depth using information related to the phase and amplitude of the received signal at the various array elements. The pattern of the individual elements and their coupling play an important role in rejecting the jamming signals while preserving the array's capability to receive information from other directions. It turns out that accurate information relating to the dielectric constants of the material surrounding the antenna array elements is also needed. Indeed, even a small variation of the dielectric constants as a function of frequency can compromise the performance of the nulling algorithm. This is due to different surface wave signal propagation speeds caused by the changing dielectric constant. As a result, the assumed linear phase variation with respect to frequency no longer holds. The phenomenon is referred to as signal dispersion and compromises the performance of the adaptive nulling algorithm. To accurately model antenna performance characteristics, it is necessary to account for the specific element pattern shape as well as coupling and dispersion effects caused by the dielectric used in place of the simulator's ideal antenna model (i.e. uncoupled, linear and isotropic elements). To do so, it is necessary to model the entire antenna structure including its detailed geometrical and material properties and to include coupling and platform curvature effects. The FEMA-PRISM code allows for the precise modeling of all

these antenna attributes and can simulate the antenna array patterns under any excitation. During the past two months the following tasks were carried out:

1. Modeled the two GPS antenna configurations using FEMA-PRISM and examined their performance around the L1 and L2 GPS bands (1227.6 MHz and/or 1575.42 MHz). The actual geometry generation, code modifications and runs were carried out by Mark Casciato who is now pursuing his Ph.D. degree at the University of Michigan. Prof. Volakis worked with Mr. Casciato in defining the analytical model, obtaining the geometry data from the USAF suppliers, data analysis, code upgrades and in planning the analysis work.
2. Wrote a 30 pp. report which provides an overview of the finite element method for antenna simulations and how the method is implemented into FEMA-PRISM. This document was also provided with five recent journal papers describing with greater detail the working principles of FEMA-PRISM and previous finite element codes which lead to the development of FEMA-PRISM. (Completed by Prof. Volakis)
3. Wrote a 30 pp. document on the basic principles of the finite element method at a level to be read by engineers unfamiliar with numerical methods and who are interested in an intelligent use of the FEMA-PRISM code. (Completed by Prof. Volakis)
4. Updated the FEMA-PRISM manual to incorporate lumped load modeling and interface with SDRC-Ideas used for meshing the antenna surface. (Done by Mr. Casciato)
5. The source code along with interfaces to MATLAB for plotting, meshing packages as well as pre- and post-processing were left with the Wright Laboratory. As is the case with all software technologies, FEMA-PRISM will need to be updated continuously to accommodate new antennas and to increase its usability to a wider community base.

Below we describe the analysis carried out for the two unique GPS antennas.

Lockheed CRPA2 Anti-Jamming GPS Array

The layout of this anti-jamming array is similar to that in Figure 1. Six anti-jamming array elements are placed around the reference element. The geometry of the anti-jam element is shown in Figure 3 based on the dimensions provided by Mr. Pellet of Lockheed. It consists of a triangular cavity which is divided in two chambers using a metallic separator with an iris placed in the middle of the separator to allow excitation of both cavity chambers. Two capacitors are used for controlling the cavity Q and thus permitting resonance at L1 and L2 bands. The location and layout of the capacitors were of sensitive nature and had to be provided through the F117 U.S. Air Force System Program Office (SPO). Unfortunately, this information was not provided till one week prior to the end of the Summer Faculty Appointment. Thus we initially proceeded to model the Lockheed CRPA2 without the capacitor elements with the intention of adding these at a later stage. Lack of the capacitor elements will cause the cavity to resonate at different frequencies and thus much of the geometry meshing and simulation can be done without a knowledge of the capacitor elements.

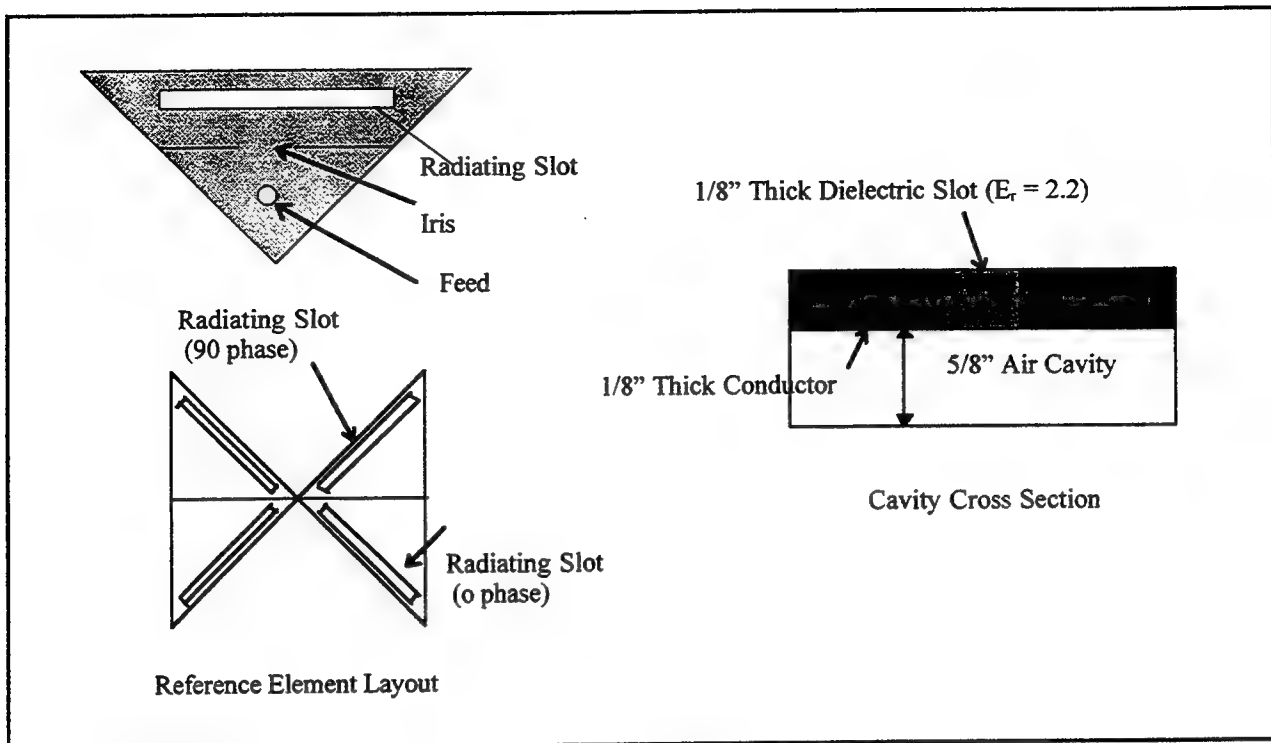


Figure 3. The Lockheed CRPA2 triangular auxiliary antenna element. Four of these constitute the reference element arranged as shown.

The cavity was fed by a coaxial cable which was modeled by a probe feed at the shown location and radiation occurred through the straight slot located near the edge of the cavity opposite to the right angle. When the cavity is at resonance, strong fields are generated within the cavity and under the slot implying maximum radiation at the corresponding frequencies. When the cavity is not at resonance, the modal cavity fields are highly attenuated before they reach the slot location. As a result, the electric field across the slot is small when the cavity is not at resonance and little or no radiation occurs at the off-resonance frequencies. The reference element consisted of four cavities arranged so that the slots form an X pattern. Since the slots are fed in quadrature with respect to the adjacent element/slot, the resulting radiation is circularly polarized whereas the excitation of the auxiliary anti-jam elements is controlled by the adaptive nulling network to generate a null in the direction of the incoming jamming signal. Thus, as a start we decided to provide a complete characterization of the single triangular cavity since the entire array, including the reference element can be modeled by repeating this element with different excitations. We will refer to the element shown in Figure 3 as the basic element of the Lockheed CRPA2.

In modeling the anti-jamming GPS antennas, the following steps were taken:

1. Translate actual model to one suitable for FEMA-PRISM simulation. This is usually an easy step and amount to certain minor simplifications. However, some difficulty was encountered due to data sensitivity issues which took some time to identify and resolve.
2. Generate a surface mesh using the FEMA-PRISM's internal meshing package or an external mesher. The latter is necessary for complex antenna configurations and for our case, we used a commercial package by SDRC-

Ideas. A translator was written for this package to reformat the data-sets. Given the surface mesh of the antenna element, FEMA-PRISM generates the volume mesh by growing prismatic elements away from the surface.

3. Post-process mesh to enforce boundary conditions and to introduce lumped elements such as resistors, capacitors, shorting pins, via holes and feed geometry. These are of course essential for the correct operation of the antenna. For the case of the Lockheed CRPA2, a special preprocessor was written to handle the conductor thickness. As shown in figure 3, the slot was formed by a 1/8" thick metal whose thickness is essential in controlling the resonance frequency. To ensure the correct operation of the PRISM code, a single slot was first modeled which was 2.75" long, 1/8" wide and was placed in a 5/8" deep rectangular cavity as in the Lockheed CRPA2 configuration. A dielectric having a relative permittivity of $\epsilon_r = 2.2$ was also placed in the slot. The results of PRISM were compared to a reference code and agreed to within 2.5% percent as shown in Figure 4. For this simulation we used 23876 non-zero edges (out of 33410) whereas the reference FEMA_BRICK code used 11948 edges. About 3500 iterations were required by FEMA_PRISM to achieve convergence.
4. Execute FEMA_PRISM code for the specified antenna geometry and excitation to obtain input impedance, near zone fields, radiation patterns, coupling effects and other parameters of interest.

The surface mesh corresponding to the antenna element in Figure 3 was generated using SDRC-Ideas and is illustrated in Figure 5. Typical results at a frequency of 1575 Mhz are displayed in Figure 6 which include two near zone plots for each of the electric field components that couple to the slot and the far zone pattern of the element. Note that the horizontal electric field component is the most dominant in the slot and is responsible for radiation. Note also that these plots were generated without the cavity separator. The convergence of the FEMA-PRISM biconjugate gradient iterative solver was extremely good for this application and required just 278 iterations for a system of about 30,000 degrees of freedom. This is close to the theoretical convergence rate of about 1/100 of the number of unknowns implying a well conditioned system. Such system characteristics are indicative of high accuracy. Also, as demonstrated by the near zone fields, the slot is excited by the cavity field.

Having successfully completed the element simulation with FEMA-PRISM we could now proceed with the array analysis, including the jamming elements with the goal of obtaining accurate patterns which include coupling effects and inherent features due to the antenna element geometry. However, given that critical information were still not available from Lockheed, we instead proceeded to model another GPS antenna candidate which employs a set of two stacked patches as described next.

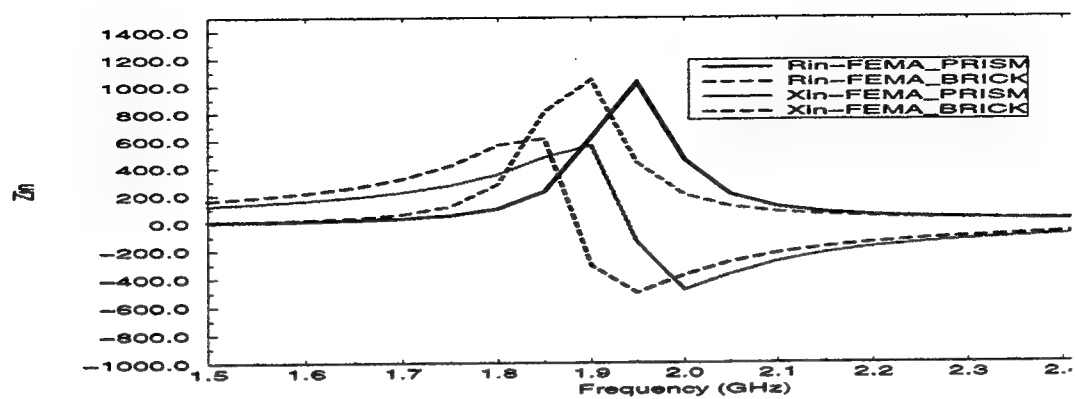


Figure 4. Slot resonance results. Comparison of FEMA-PRISM and FEMA-BRICK

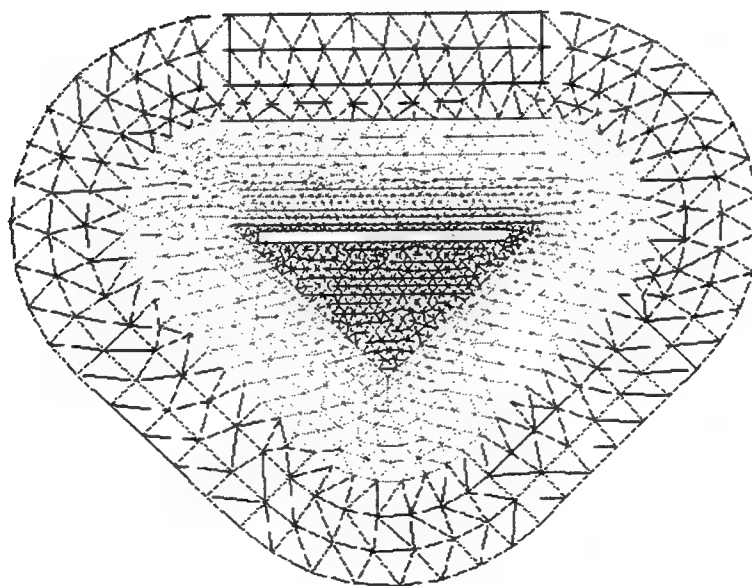


Figure 5. Surface Mesh of the Lockheed CRPA2 basic element

TRIANGULAR CAVITY SLOT

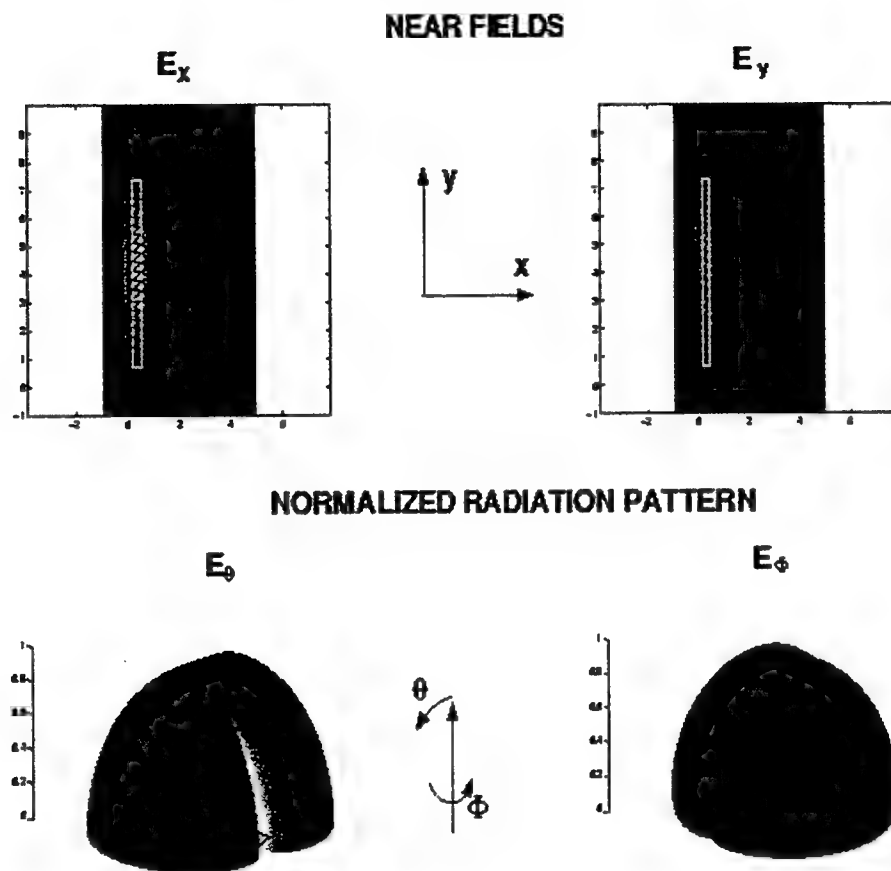


Figure 6. Results of FEMA-PRISM for the Lockheed CRPA2. Near zone fields and radiation pattern.

Ball Communications Dual Patch Anti-Jamming GPS Array

The Ball Communications dual patch array is a similar configuration to Figure 1. The Ball elements however consist of a stacked patch configuration as shown in Figure 7. In this configuration the smaller top patch is for the L1 (1574.42 Mhz) frequency with the inner-cavity patch for L2 (1227.6 MHz). A coaxial cable, routed through a hole in the L2 patch, drives the L1 patch with the L2 patch being parasitic. Again the coax is modeled as a current probe feed.

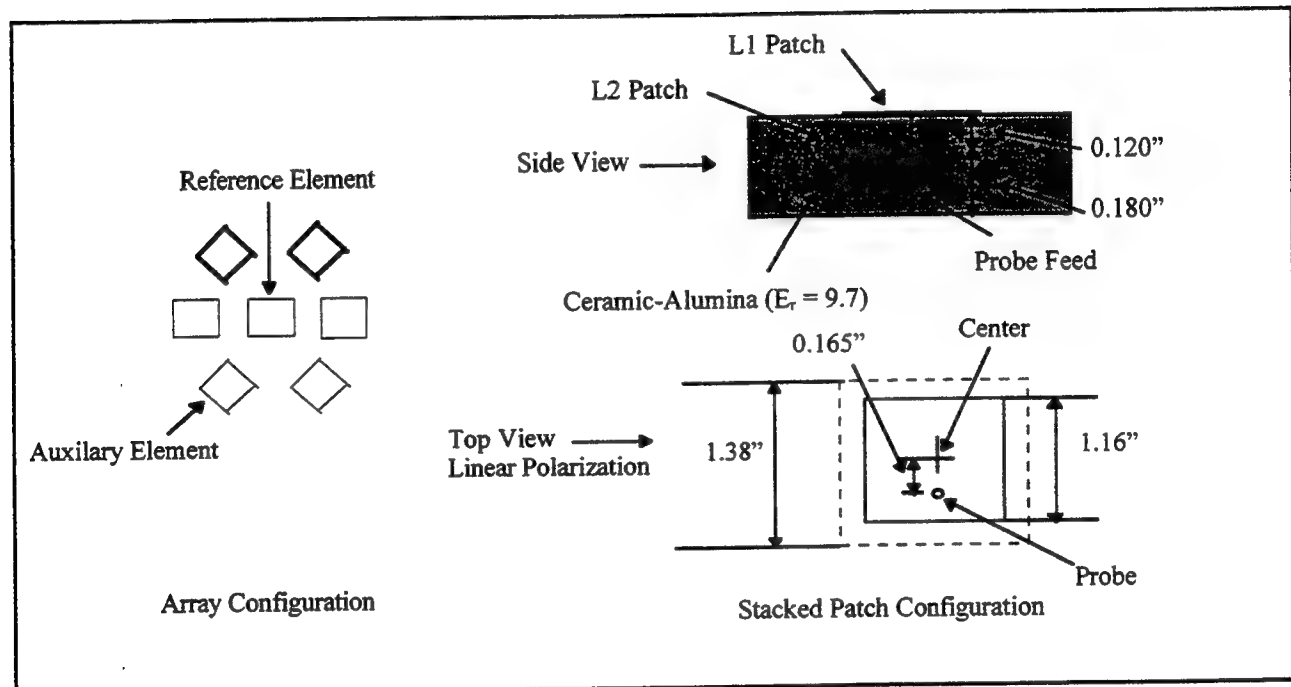


Figure 7. Ball Dual Patch Antenna. Array and Stacked Patch Configuration.

The array layout is also shown in Figure 7. As was the case with the Lockheed antenna detailed information on the Ball antenna was not available at this time. Due to this, a stacked patch configuration as shown in Figure 7 was used[13]. Note that the probe's offset location from center guarantees linear polarization. A second probe, with a 90° phase shift from the one shown, placed right/left of patch center will create circular polarization. The following task were carried out in connection with the Ball array.

1. To account for the larger L2 patch and hole, code modifications were made to permit a varied patch and hole size in specified layers. To validate these changes a single reference patch was generated and again compared to FEMA-BRICK data[13]. The results showed the antenna element resonances at the proper L1 and L2 locations as illustrated in Figure 8. However, in comparison to the FEMA-BRICK data[3] a difference in absolute impedance levels was observed. A possible explanation for this difference is that while the hole was modeled as perfectly square in FEMA-BRICK, for the FEMA-PRISM model the central patch along with the patch itself had trimmed corners. Figures 9a,c show the normal directed fields directly under each patch overlaid with the patch outline. As can be seen in Figure 9c the fields under the L2 patch show perturbations due to the trimmed

corners which will have an effect on the impedance levels. Also absolute impedance levels are difficult to predict. Figures 9b,d show the respective far-field patterns of the L1 and L2 frequencies. For the results in Figure 9, the linear system was approximately 71,000 unknowns and convergence was achieved in 1800-4000 iterations depending on the frequency of calculation. Typically, the convergence is slower at the resonance frequency because of the higher field singularities. The runs were executed on an SGI Indigo 2 and the CPU time per iterations was 2 seconds.

2. Generate a surface mesh using SDRC-Ideas of the full array, time permitting
3. Execute FEMA_PRISM for the full array geometry including mutual coupling factors, time permitting. Due to the limited information available the stacked patch configuration from [13] will be used to model the Ball array. This will be a proof of concept showing FEMA_PRISM's ability to model all array interactions including nulling produced by the auxiliary elements.

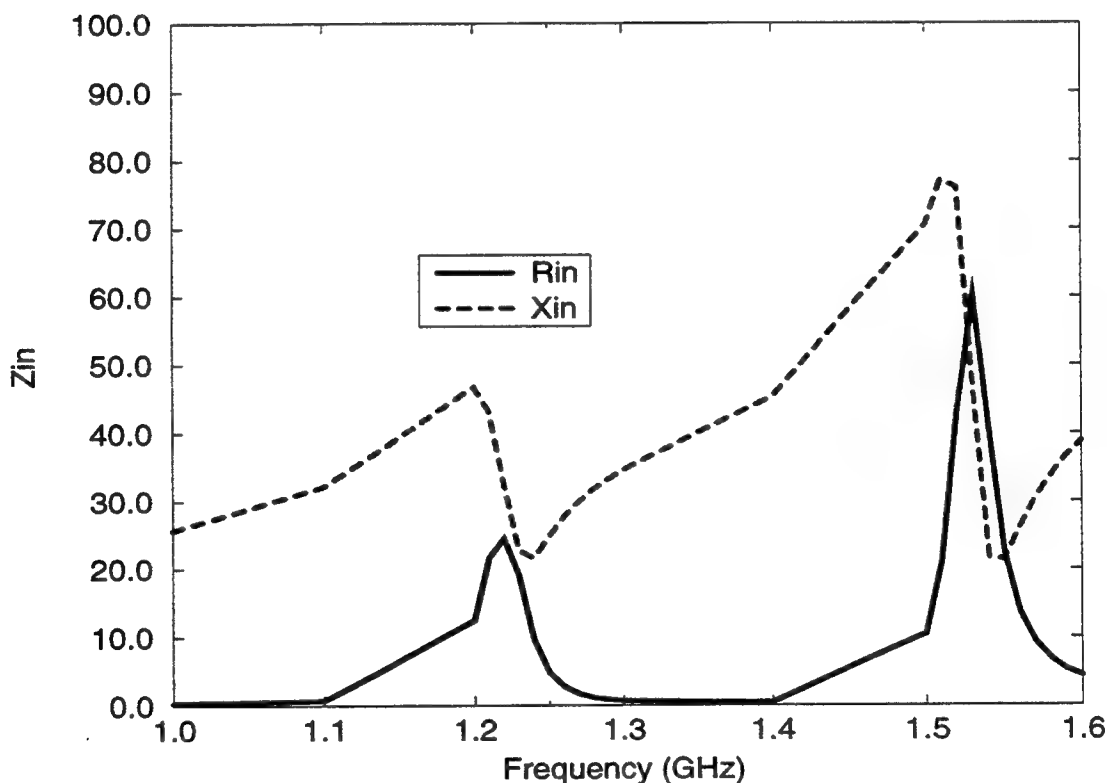


Figure 8. Input impedance of the dual patch Ball Communications element as a function of frequency. Note that the spikes in the plot indicate the frequencies at which the patch is resonant.

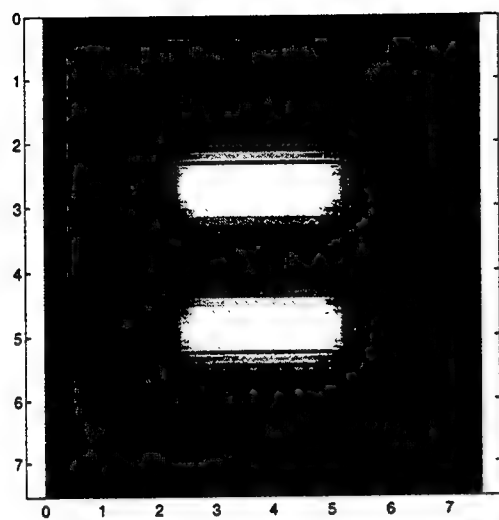


Figure 9a. Substrate Fields at L1 band.

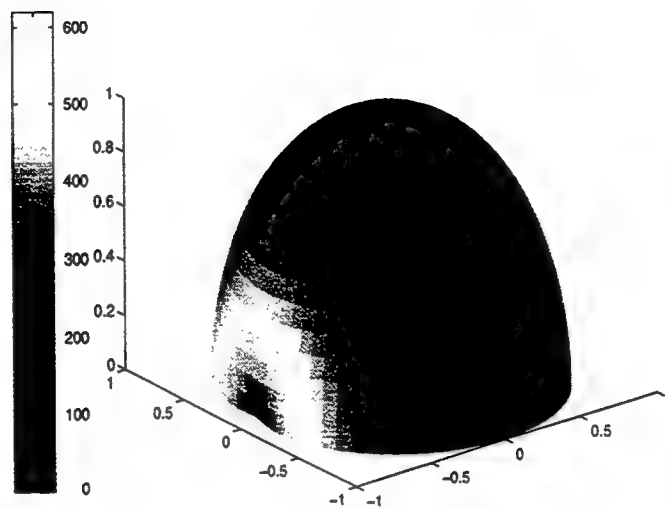


Figure 9b. Far-Field Pattern at L1 band.

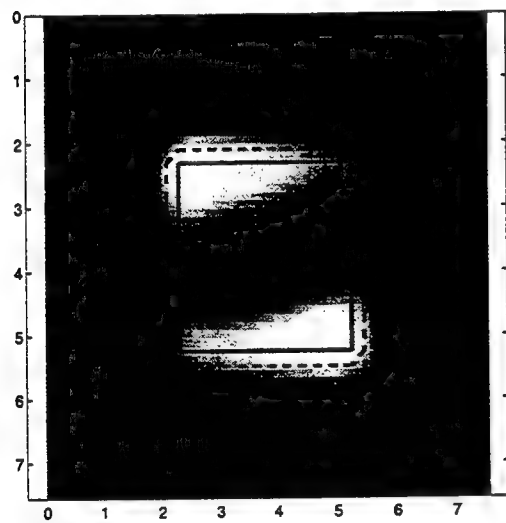


Figure 9c. Substrate Fields at the L2 band.

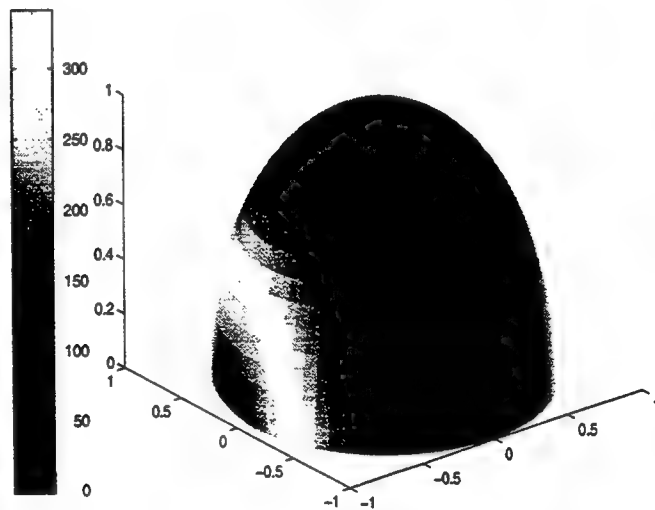


Figure 9d. Far-Field Pattern at the L2 band

FEMA PRISM Code Improvements and Upgrades.

Several changes were made to FEMA_PRISM and supporting codes to allow modeling of the anti-jam arrays. In addition many of these changes are applicable to other problems and enhance the abilities of FEMA_PRISM to handle future antenna configurations. All code modifications were made with the user in mind. With few exceptions, most changes are transparent to the user and are simply a question of modifying the FEMA_PRISM input file, MainInput. Also, where necessary, the changes were validated with simple examples and compared to reference data.

1) PreProcessing:

The ability to process CAD files is essential for a code to analyze various antenna platforms. CAD packages such as SDRC-Ideas, Hypermesh, etc., allow the user to easily generate various surface meshes of high quality. Output files contain information such as node location, connectivity, material identifiers, conducting pin location, etc. With this in mind a preprocessor was written to operate on an SDRC-Ideas universal file and convert it to a FEMA_PRISM compatible mesh. Note that other CAD files use similar methodology to describe mesh information. As other CAD packages are evaluated it is simply a matter of modifying the preprocessor to read in the new file format. Everything else is taken care of automatically. A sample preprocessor input file with comments is shown below.

```
!Input Universal File Name.
slele1.unv
!Output FEMA_PRISM File Name.
test1.mesh
1      !Node Convention (1)LHN,(2)RHN.
-45    !Object Rotation (Degrees).
2      !Cad File Units.
        !Meters(1), cm(2), mm(3), Feet(4), inches(5)
0      !Iris for Slot Dual Cavity: no(0), yes(1)
```

As can be seen the preprocessor reads in an Ideas universal file, and outputs a FEMA_PRISM compatible mesh. The user specifies the universal input file node convention (FEMA_PRISM internally uses RHN convention); degrees to rotate the mesh if desired; and the universal file units (FEMA_PRISM uses cm). The last option is experimental to read in iris information for the Lockheed CRPA2, and should be left at 0 for most applications.

For a final visual check, the preprocessor writes out the FEMA_PRISM mesh file and also generates it's own universal file, chkslt.unv, with each material parameter given a separate color so that the user can visually check the final mesh and mesh parameters in the CAD package. Note that in the universal file nodes are written first, followed by the connectivity tables. Next, the conducting (grouped in Ideas by the user) and the iris pin nodes are listed, if necessary. The following material convention was used:

<u>Material #</u>	<u>Material Type</u>
1	air
2	absorber
3	dielectric under conductor(surface patch)
4	dielectric (exposed portion of dielectric)

2) Internal Rectangular mesh generator:

FEMA_PRISM contains an internal mesh generator which generates a simple rectangular mesh. The option was added in the MainInput file to generate a rectangular slot instead of a patch.

3) Conducting Pins:

An important addition to FEMA_PRISM is the ability to define conducting edges or pins in the cavity volume. In the case of FEMA_PRISM, these are vertically directed pins i.e., pins normal to the mesh surface. The surface position of these pins are identified in the Ideas universal file and automatically added to the FEMA_PRISM input surface mesh file. Along with inner-cavity conducting patches, these pins allow the user to define a finite thickness of the conducting patch. For resonant cavities, a finite conductor thickness as opposed to a zero thickness conductor, makes for a more accurate and stable numerical solution. These pins are added to layers in the cavity and must therefore conform to the user defined layers in the cavity. To add conducting pins in any layer, the user in the FEMA_PRISM input file, specifies the number of layers that will contain conducting pins, and then lists the pin layers.

4) Conducting Layers:

The ability to specify conducting layers or inner-cavity patches is another important addition to FEMA_PRISM. Coupled with conducting pins, these allow the user to define a conductor thickness. In addition, conducting layers allow for modeling stacked patch antenna configurations. As with conducting pins, the user specifies in the input file the number of conducting layers and lists the appropriate layers. At the top of each layer, a duplicate of the surface conducting patch is inserted. Again the specified layers must conform to the user defined cavity layers.

5) Expanded Layers/Holes:

With the ability to expand conducting patches in any layer and to insert a hole in these patches FEMA_PRISM can now analyze stacked patch antenna configurations. The expanded layer must be one of the conducting layers defined previously. In the FEMA_PRISM input file, the user simply defines the number of segments to expand the patch. For each segment, the edges radiating from the conducting patch along with the edges connecting these radiating edges are defined as conducting. For patch holes, the layer containing the hole must be one of the conducting layers defined previously. The user specifies the layer containing the hole, the number of holes in that layer, the hole locations (i.e., node where the holes will be centered is usually the same location as the feed probe), and the number of segments/elements to expand the square hole. All holes in a given layer will be expanded by the same amount. For the first expanded segment, edges radiating from the specified nodes will be defined as non-conducting. For each additional segment the hole is expanded, non-conducting hole edges defined in the previous iteration will be

connected by non-conducting edges, and then edges radiating from these new non-conducting edges will be defined as non-conducting.

6) Iris:

For the Lockheed CRPA2 triangular cavity the option to add a conducting wall to the cavity was added. The surface nodes where the wall is placed is identified again in the Ideas universal file and transferred to the FEMA_PRISM mesh file. The wall is then automatically grown between the specified nodes. If a vertical iris is desired in the wall the opening must be at least 2 segments wide or the wall will extend the breadth of the cavity with no opening.

7) QMR option:

In addition to the currently available BiCG iterative solver, the option of a QMR solver was added. This solver converges quasi-monotonically as opposed to the oscillating convergence of BiCG. While the QMR solver was fully tested for accuracy, due to time limitations, it wasn't optimized for speed and further work is needed. The user specifies in the input file whether BiCG or QMR is preferred.

8) Diagonal Preconditioner:

A diagonal preconditioner was added to FEMA_PRISM to increase the rate of convergence. It was found that while a significant improvement was seen when QMR was specified, BiCG reacted unfavorably to this type of preconditioning. Consequently, preconditioning is only implemented internally when QMR was specified. Note that the pre-conditioned QMR converges at approximately the same rate as the unconditioned BiCG.

Conclusions and Future Directions

The most critical aspect of this Tour has been the identification of a situation, where a new software technology can have a major impact on our ability to develop real world anti-jamming array simulations without resorting to extensive measured databases that are likely to be impractical and unsuitable for array optimization and improvements. Thus, it is important to note that the antenna characterization tasks were carried out by FEMA-PRISM methodically and without difficulty. This clearly demonstrates the code's capability to handle these GPS communication antennas as well as other configurations used on GPS-guided munitions. The latter are mounted on doubly curved platforms and a unique feature of FEMA-PRISM is its capability to readily handle arrays on nonplanar surfaces and to account for coupling and element interactions in the presence of complex materials. Presently, the array simulation models used in conjunction with the Wavefront Simulator (AWFS) facility are based on "ideal" isotropic sources, and thus the FEMA-PRISM code can become an invaluable and possibly necessary addition to the AWFS. It can provide pattern data for the actual array configuration and this data can be directly fed to the AWFS for real-time evaluation of the array performance under various jamming conditions. That is, the array analysis software can be installed on a dedicated computational workstation and integrated with the AWFS for a performance evaluation under conditions which are nearly identical to the real-world situation. The risk for the success of this hardware-in-the-loop simulator is indeed very low and the cost is limited to the specialization of the software for such an application. Before, proceeding with the implementation of such a computational workstation, it is necessary

to make further utilization of the code for various arrays
exercize the code in connection with a variety anti-jamming scenarios
make extensive comparisons with measurements
simulate prototype situations of the integrated workstation platform
develop suitable I/O interface to the code for ease of use as part of the AWFS

These tasks are, of course, quite substantial since they outline a project that will be of major benefit to the Air Force in terms of simulation accuracy and cost-reduction. The Summer Faculty Research Continuation Grant appears to be a most prudent way to continue toward the completion of these tasks. Of course, the realization of these tasks and goals cannot be achieved without the commitment of substantially more resources.

Appendix: UNIVERSITY OF MICHIGAN FINITE ELEMENT CODES

Over the past 10 years we have witnessed an increasing reliance on computational methods for the characterization of electromagnetic problems. Although traditional integral equation methods continue to be used for many applications, one can safely state that in recent years the greatest progress in computational electromagnetics has been in the development and application of partial differential equation(PDE) methods such as the finite difference-time domain and finite element (FEM) methods[3], including hybridizations of these with integral equation and high frequency techniques. The major reasons for the increasing reliance on PDE methods stem from their inherent geometrical adaptability, low $O(N)$ memory demand and their capability to model heterogeneous (isotropic or anisotropic) geometries. These attributes are essential in developing general-purpose codes for electromagnetic analysis/design, including antenna design and characterization. Other attributes of the finite element method are:

- input data to FEM software can be extracted directly from commercially available (i.e. well-supported) solid modeling packages which run on all popular workstation platforms and are well documented. This is particularly important to problems in antenna analysis and design, where a high degree of geometrical fidelity must be maintained(see Figure 1 for examples of antenna and circuit meshes).
- FEM is totally insensitive to the material composition of the radiating or scattering structure. Also, resistive/material and impedance boundary conditions are readily implemented in a modular fashion.
- Being a near-neighbor method, new "physics" can be added to the FEM codes without a need to alter the original code structure. Neither the moment method nor the finite difference method share this feature.
- Being a frequency domain method, the FEM is ideal for antenna analysis and design purposes. Established hybridizations of the FEM with moment method and ray methods provide an added advantage by delivering the most adaptable and efficient code when compared to other approaches.

- Advances in mesh termination schemes have relaxed accuracy compromises with that aspect of the method. Also, the FEM can benefit from recent fast integral equation algorithms. One may therefore think of the FEM as the core method for treating the heterogeneous volumetric structures including feeds and loads
- Fine geometrical details such as those found in spiral antennas and the feeding structure can be modeled without sacrificing geometrical fidelity and CPU efficiency.
- The immediate output of FEM codes is the near zone fields which can be readily visualized (superimposed with the geometry) and further processed using commercially available tools. Moment method codes do not share this important feature. This inherent aspect of the FEM codes allows for extraction and visualization of many different parameters as needed by the user.

Several finite element codes have been developed at the Univ. of Michigan for the analysis and design of printed antenna configurations. Typically, the printed antenna configuration is assumed to be recessed in some metallic or coated platform and the various codes differ in the element used for the tessellation of the antenna, the type of platform assumed in the analysis (planar, cylindrical or doubly curved) and the closure condition employed for terminating the finite element mesh. The following codes are available for antenna radiation and scattering analysis.

FEMA-CYL: This code is specialized to cavity-backed antennas recessed in a metallic cylindrical platform. The finite element-boundary integral method is employed for truncating the FEM mesh and to simplify the geometry I/O, cylindrical edge-based shell-elements are used for the discretization of the cavity and the aperture. As a result, the mesh generation is done automatically and this process does not require interaction with the user. Multiple patches and patch arrays (consisting of individual cavity-backed elements or elements on a single substrate) can be considered but their geometry is restricted by the generated grid. This situation is similar to finite difference codes, where the model's boundary is modified to fit the geometry. Consequently, although this code is user-friendly, it is not appropriate for computing the radiation parameters of circular patches or spirals. However, scattering parameters are not as sensitive to the minor geometrical modifications and therefore FEMA-CYL can still be used for computing the scattering of non-rectangular printed antenna configurations and cavities.

FEMA-CYL has been validated for antenna and scattering applications. For fast execution, it employs the FFT for performing the matrix-vector products in the iterative solver. Both, the biconjugate gradient and QMR solvers are available as options and by virtue of the FFT, the code's computational and memory requirements remain at $O(N)$. The system solution generates the fields within and on the surface of the antenna cavity. For the radiation and scattering pattern calculations only the surface electric fields are used to generate equivalent magnetic currents which are then used in the radiation integral with the appropriate platform Green's function to generate the antenna parameters. For antenna excitation, a probe or a coaxial cable feed model is employed and in the case of scattering the excitation become the aperture fields due to the incoming field. When a probe feed model is used, the input impedance is obtained by dividing the computed field at the probe location with the probe current. When a coaxial

feed is employed, the excitation is introduced by setting the fields coinciding with the edge elements bordering the opening of the coax cable equal $\Delta V/\Delta L$, where ΔV is the given potential between the outer and inner surface of the cable and the $\Delta L=b-a$, where b and a correspond to the outer and inner radius of the cable. The center conductor is modeled by setting to zero the mesh fields at location of the vertical conductor which connects the antenna to the cable.

FEMA-CYL can also be used to include the effect of superstrate materials which may extend over the antenna platform. To avoid use of complicated Green's functions the mesh may be extended a small distance over the cavity's aperture and an absorbing boundary condition is employed for truncating the FEM mesh. With this type of mesh truncation, the entire system is sparse and thus the memory and CPU requirement are again $O(N)$ without a need to make use of the FFT. We have found that placement of the mesh at a distance of 0.3λ away from the coating's surface is sufficient to obtain the radiation pattern with reasonable accuracy. As before, the pattern is computed by introducing equivalent electric and magnetic currents placed at the surface of the dielectric coating and then integrated using the radiation integral.

FEMA-BRICK: This code is a special version of the FEMA-CYL code and is applicable only for cavity-backed antennas recessed in planar platforms. It is also referred to as XBRICK in the community and employs brick elements for the discretization of the cavity volume. Thus, as was also the case with FEMA-CYL, the geometry of the antenna must be modified to fit the grid. Because of the simplicity of the Green's function and the use of the FFT on the boundary integral, the XBRICK is the fastest code among all mentioned here. The code is also provided by Wright Laboratory with a graphical interface to enter the geometry and specify the material layers, loads and printed elements.

FEMA-TETRA: This code employs the same FEM formulation as FEMA-CYL and FEMA-BRICK for computing the antenna and scattering parameters of printed antennas recessed in a metallic platforms. In contrast to FEMA-CYL, this code employs tetrahedral elements for modeling the radiating structure and bricks or shells for modeling the feed. As a result, it incorporates maximum geometrical adaptability and has already been employed for the analysis of antenna+feed configurations beyond the capabilities of moment method codes. The FFT can also be invoked for carrying out the matrix-vector products associated with the boundary integral portion of the system in the iterative, thus maintaining an $O(N)$ memory requirement for the entire system. Its accuracy has already been demonstrated for a variety of antennas, including stacked patches and finite arrays in ground planes, printed and slot spirals and aperture antennas. In the case of aperture fed patch configurations, the mesh generation is divided into the regions above and below the feeding slot. The mesh in each region can employ different elements and this is important in avoiding meshing bottlenecks due to the narrowness of the slot. The mathematical connection between the meshes are then mathematically connected by enforcing electric field continuity across the slot. However, since the mesh elements in each regions are different, this condition is implemented by first relating the fields in the slot

to the potential existing across the slot. Using this procedure, there is no need to have coincidence of the edges bordering the slot and is the main reason which allows use of different elements in the regions above and below the slot aperture. The FEMA-TETRA code is interfaced with commercial solid-modeling packages such as SDRC I-DEAS, Hypermesh and PATRAN to enter the geometry and generate the volume mesh. Using such geometry modeling packages and owed to the adaptability of tetrahedrals, nearly any printed antenna configuration can be modeled and analyzed using FEMA-TETRA. As can be expected, because FEMA-TETRA is not specialized to any antenna configuration all geometrical data must be entered through the solid-modeling package and this may be a time consuming task for inexperienced users.

FEMA-PRISM: To avoid the time consuming task of volume mesh generation without compromising geometrical adaptability, this code provides an attractive compromise by making use of prismatic elements. Prisms have triangular faces at their top and bottom faces and as a result they can model any patch configuration with sufficient geometrical fidelity. Typically, since the substrate/substrate is of constant thickness, once the triangular grid is constructed, the prismatic volume mesh can be generated by growing the elements above and below the surface of the printed antenna. Thus use of prismatic elements reduces the mesh generation task from a three-dimensional to a two-dimensional task. That is, the user needs to only generate the surface mesh in the aperture containing the printed antenna and the volume mesh is then generated automatically by specifying the substrate/substrate thickness, aperture size and information relating to the closure condition. FEMA-PRISM makes use of an ABC or artificial absorbers for truncating the mesh at some distance from the aperture surface. As in the other codes, the radiation and scattering patterns are computed using the aperture fields and the corresponding equivalent magnetic currents in the radiation integral with the free space Green's function as the propagator. Also, since a non-integral closure condition is used for truncating the mesh, the antenna platform can be planar or non-planar (doubly curved). It is, of course, assumed that the generated results will not include effects due to platform interactions and to include these it is more appropriate to interface the FEM codes with high frequency analysis codes[11]

References

1. T. Ozdemir and J.L. Volakis, "Users manual for FEMA-PRISM, " Univ of Michigan Radiation Lab. Techn. Report 031307-6-T, March 1996. 15pp.
2. T. Ozdemir and J. L. Volakis, "Triangular prisms for edge based vector finite element antenna analysis," U-M Radiation Laboratory report 031307-4-T, March 1995, 20 pp
3. J. L. Volakis, A. Chatterjee and L.C. Kempel, "A review of the finite element method for three dimensional scattering," *J. Optical Society of America A*, April 1994, pp. 1422-1433,.
4. J.Gong and J.L. Volakis, A. Woo and H. Wang, "A hybrid finite element method-boundary integral method for the analysis of cavity-backed antennas of arbitrary shape," *IEEE Trans. Antennas Propagat.*, 42, Sept. 1994, pp. 1233-1242.

5. J.L. Volakis, J. Gong and T. Ozdemir, "Applications of the FEM to conformal antennas," chapter to appear in book edited by Itoh, Silvester and Pelosi, Wiley, 1996.
6. J.L. Volakis, J.Gong and A. Alexanian "A finite element boundary integral method for antenna RCS analysis," *Electromagnetics*, **14**, No.1, 1994, pp. 63-85,.
7. J.M. Jin and J. L. Volakis, "A hybrid finite element method for scattering and radiation from microstrip patch antennas and arrays residing in a cavity," *IEEE Trans. Antennas and Propagat.*, **39**, Nov. 1991, pp. 1598-1604,
8. L.C. Kempel, J.L. Volakis and R. Sliva, "Radiation by cavity-backed antennas on a circular cylinder," *IEE Proceedings-H*, , 1995, pp. 233-239
10. T. Ozdemir, J.L. Volakis, et.al. "Finite element scattering and radiation analysis using prismatic meshes and artificial absorbers for conformal domain truncation," 1996 ACES conference Proceeding, pp. 1174 - 1180.
11. T. Ozdemir, M.W. Nurnberger, J.L. Volakis, R. Kipp and J. Berrie, "A hybridization of finite element and high frequency methods for pattern prediction of antennas on aircraft structures," *IEEE Antennas and Propagat. Soc. Mag.*, June 1996.
12. Dana Howell and Bruce Rahn, " Antenna WaveFront Simulator(AWFS)," Wright Laboratory(WL/AAI-1), WPAFB, OH 45433
13. S.W. Schneider, "On the Analysis of a Stacked Patch Antenna", Wright Laboratory Memorandum.

AN EXPERIMENTAL AND COMPUTATIONAL ANALYSIS OF THE INFLUENCE OF A
TRANSONIC COMPRESSOR ROTOR ON UPSTREAM INLET GUIDE VANE WAKE
CHARACTERISTICS

J. Mitch Wolff
Assistant Professor

Department of Mechanical and Materials Engineering

Wright State University
3640 Colonel Glenn Highway
Dayton, OH 45435-0001
Ph: (513) 873-5040
Fax: (513) 873-5009

Final Report for:
Summer Faculty Research Program
Summer Graduate Student Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright-Patterson Air Force Base, Dayton, OH

September 1996

AN EXPERIMENTAL AND COMPUTATIONAL ANALYSIS OF THE INFLUENCE OF A
TRANSONIC COMPRESSOR ROTOR ON UPSTREAM INLET GUIDE VANE WAKE
CHARACTERISTICS

J. Mitch Wolff
Assistant Professor

Department of Mechanical and Materials Engineering
Wright State University

Abstract

Gas turbines are a vital energy source for both industrial and military applications. Recent research has focused on identifying the unsteady flow mechanisms inherent in gas turbines. A progress report is given of a research study utilizing both a computational analysis and experimental tests on the unsteady aerodynamic effect of a downstream transonic compressor rotor on the upstream inlet guide vane/stator wakes. This study should show that the wakes coming off of the upstream stators are significantly affected by the unsteady action of the transonic rotor with its inherent shocks. Currently, design codes do not consider the unsteady aerodynamic effects of the rotor/stator interaction. The purpose of this current study is: (1) experimentally determine if the inlet guide vane surface pressures are changed and gain insight into the driving unsteady aerodynamic processes (2) computationally model the experimental configuration using a nonlinear unsteady multi blade row Navier-Stokes code, identify model weaknesses and make improvements to the unsteady turbomachinery computational code. This report describes the experimental setup and the computational method selected.

AN EXPERIMENTAL AND COMPUTATIONAL ANALYSIS OF THE INFLUENCE OF A TRANSONIC COMPRESSOR ROTOR ON UPSTREAM INLET GUIDE VANE WAKE CHARACTERISTICS

J. Mitch Wolff

Introduction

The performance of aircraft gas turbine engines must undergo continual improvements to meet durability, reduced noise levels, and increased performance of future flight systems. This need has motivated engineers to assess the effects of flow unsteadiness on the time-mean aerodynamics in axial-flow turbomachines. Of special interest in the overall engine performance is the compression system; therefore improved understanding of axial-compressor flow physics must be continuously developed to meet required engine performance standards. In particular, recent research has focused on identifying and controlling the unsteady flow mechanisms associated with the interaction between adjacent blade rows in gas turbines.

The two principal types of such interaction are usually referred to as potential-flow and wake interactions (Verdon, 1992). Potential-flow interaction results from the variations in the velocity potential or pressure fields associated with the blades of a neighboring row and their effect upon the blades of a given row moving at a different rotational speed. This type of interaction is of serious concern when the axial spacing between adjacent blade rows is small or the flow Mach number is high. Wake interaction is the effect upon the flow through a downstream blade row of the vortical and entropic wakes shed by one or more upstream rows.

Recently, computational work has been initiated to develop nonlinear, time-accurate, inviscid (Euler) and viscous (Navier-Stokes) solution techniques for unsteady flows through isolated and aerodynamically coupled blade rows (see Verdon, 1992 for a review). For coupled systems of rotating and stationary blade rows, the relative motions between adjacent rows give rise to unsteady aerodynamic excitations which can excite blade vibrations, generate discrete-tone noise, and degrade aerodynamic efficiency. Two categories of

numerical procedures have recently been developed for determining the effects of relative motion between adjacent blade rows. In the first category of numerical procedures, incoming wakes are specified at the inlet of isolated blade rows (Suddhoo et al. 1991). In these methods the wakes are usually assumed to be parallel, with uniform pressure and prescribed total enthalpy and/or velocity variations. In the second category of numerical analyses, both blade rows are modeled and the relative position of one blade row is varied to simulate blade motion (Giles, 1988; Rai, 1989; Rao and Delaney, 1990). The computational models developed have been primarily applied to turbine configurations because the favorable pressure gradient greatly simplifies the aerodynamic analysis.

Some joint computational and experimental investigations have been made into vane/blade interactions. Again these investigations have been primarily on turbine configurations with the experimental data acquired on generally two different types of experimental rigs. First, large scale turbine rigs are used to simplify the experimental investigation. The large scale of the rig is a distinct advantage because it permits the use of extensive instrumentation on both the stationary and rotating airfoils. The large scale also has the advantage of giving Reynolds numbers which are typical of high pressure turbines at nominal model running conditions (Dring et al., 1982), but it can not simulate transonic flow phenomena. Full-scale transonic turbines are being tested using blow down or shock tube facilities (Dunn et al., 1992). These facilities are beneficial for testing actual hardware components with research issues being miniature measurement techniques and short test duration times.

This report is a progress report of a research study to obtain and compare data from both a computational analysis and experimental tests on the unsteady aerodynamic effect of a downstream transonic compressor rotor on the upstream guide vane/stator wakes. It is expected that the wakes coming off of the upstream stators are significantly affected by the unsteady action of the transonic rotor with its inherent shocks (bow and oblique passage shocks). Current compressor design codes do not consider the unsteady aerodynamic effects of the rotor/stator interaction. The purpose of this current study is (1) experimentally determine if the Inlet Guide Vane (IGV) surface pressures are changed and gain insight into the driving unsteady aerodynamic processes in a full scale compressor (2) computationally model the experimental

configuration using a nonlinear unsteady multi blade row Navier-Stokes code, identify model weaknesses and make improvements to the unsteady turbomachinery computational code.

Compressor Aero Research Laboratory (CARL)

The experimental portion of this research investigation will be completed in the Compressor Aero Research Laboratory (CARL). The facility consists of an open or closed loop (currently open) tunnel system with an upstream venturi flow meter. The test compressor is driven by a 2,000 hp electric motor with a variable speed range of 6,000 to 21,500 RPM. Other parametric ranges of the system are a corrected mass flow rate of 10 to 60 lb/sec, an inlet total pressure of 6 to 15 psia, and a discharge maximum temperature of 600° Fahrenheit. The operation of the rig is set up to have a manually controlled speed, throttle, and variable geometry. The measurement capabilities of the data acquisition system is as follows; 160 channels for pressure, 100 channels for temperature, 10 channels for rotor telemetry, and 14 channels for high response pressure measurement (250 Khz).

The current research compressor rig, Figure 1 (Cheatham, 1991), was designed by CARL personnel and manufactured by Pratt and Whitney Aircraft Engines. The primary intent for this research compressor is for a Stage Matching Investigation (SMI) characterizing overall compressor performance. Therefore, a single stage core compressor consisting of a rotor and stator with 33 and 49 airfoil blades respectively is used. The outer diameter for the rotor and stator is 19 inches. The SMI's core compressor design results in a transonic rotor. To study the effect of different upstream stages, an IGV assembly is placed upstream of the core compressor. The inlet guide vanes were designed by Pratt and Whitney with the purpose of creating a propagating wake consistent with a modern technology, highly loaded, low aspect ratio stator. They do not turn the flow as would a normal assembly of IGV's. The IGV's have a unique feature, in order to achieve a uniform wake, of a constant solidity along the span. With this design there are several ways in which to modify the wake profile being generated. The first is to change the number of IGV blades in the upstream passage. A split ring assembly is used for installation and three different configurations of IGV blades can be utilized (12, 24 and 40). Also, it is possible to vary the axial spacing between the rotor and the IGV

blades. Three different spacings are possible 0.299, 0.776, and 1.791 inches from the IGV trailing edge to the rotor hub leading edge.

A unique feature of the current test configuration is the Circumferential Traverse Assembly (CTA). The CTA is designed to be installed between the IGV's and the rotor. It has three total pressure probe rakes located at three different axial positions. The axial positions correspond to different possible spacings between the IGV's and rotor. The total pressure probes can then be traversed in the circumferential direction either manually or automatically across a complete flow passage. Therefore, using the CTA the flow characteristics of the IGV blade wake can be experimentally obtained. The rotor leading edge is 2.0 inches axially downstream from the closest CTA pressure rake. This distance is designed to enable CTA pressures to be obtained without any unsteady interaction.

IGV Surface Pressure Measurements

After all of the different IGV configurations have been characterized, the CTA will be removed and the compressor stage performance utilizing the different IGV configurations will be obtained. Since the unsteady flow mechanisms associated with the interaction between adjacent blade rows is unknown, it was proposed that miniature high response surface pressure transducers be mounted on the IGV blades. These transducers will be used to obtain information pertaining to the unsteady flow mechanisms. Mounting pressure transducers directly on the IGV blade surfaces provides a relatively inexpensive means of gaining insight into the magnitude of the unsteady interaction. Eventually, the measurement of the flow characteristics between the IGV's and the rotor will give even more detail of the flow physics but these measurements are quite expensive and time consuming to obtain. Currently several means of making non-intrusive measurements in this area are being pursued by CARL personnel.

Since the purpose of the IGV surface pressure transducers is to identify the unsteady vane/blade interaction, their location was specified to optimize the acquiring of this information. Figure 2 shows the chordwise and spanwise locations of the 20 miniature high response pressure transducers (10 on each side of the blade). The chordwise pressure resolution is greatest near the trailing edge of the IGV blade because this

is the region of unsteady interaction. No unsteady response is expected over the first half of the blade chord. The spanwise locations were selected to enable information about the two dimensionality of the flow field to be determined. Also, the inlet relative Mach number varies from approximately 0.9 to 1.3, hub to tip, respectively, so stronger shock effects can be studied by placing the transducers closer to the tip. The 24 blade IGV configuration will be instrumented. This will leave the option of testing the 12 blade configuration by removing half of the blades. This is necessary since the rig is constructed with the IGV blades secured in a separate housing which must be modified to get the leads out.

The pressure transducer selected is the LQ-125 made by Kulite. It has a pressure operating range of 0-25 psia and a natural frequency of 300 KHz. The compensated temperature range is 30-130 degrees Fahrenheit with a one percent thermal zero and sensitivity shift. A chip-on technique will be utilized to ensure minimal disturbance of the flow. Therefore, Kulite will assemble the LQ-125 series transducer and temperature compensation module on the actual IGV blades. The blade pass frequency for the SMI rotor is 7.5 KHz. The pressure transducers should be able to resolve 10 times the blade pass frequency to ensure the unsteady phenomena is captured (i.e. 75 KHz frequency response is needed). Typically 50% of the natural frequency is considered usable for Kulite pressure measurements, therefore this pressure transducer meets the frequency requirement. The transducers have been ordered from Kulite with delivery of the two instrumented IGV blades expected in mid-December.

Computational Model

Recent developmental efforts in the turbine engine propulsion arena have focused on the computation of unsteady flows. This trend will continue because virtually all components of gas turbines (e.g. compressors and turbines) produce unsteady fluid flow phenomena. Advances in the field of unsteady turbomachinery computational models have reached the point where 2-D and 3-D unsteady Reynolds-averaged Navier-Stokes models are being developed and validated. A major problem confronting validation exercises for unsteady turbomachinery flows is the lack of experimental results in which a high level of confidence exists. The first part of this report described research work that will result in unsteady vane/blade

experimental data for a high speed transonic compressor. The second part of this report will describe the unsteady 2-D vane/blade computational model utilized for comparison with the experimental data.

For several years, NASA Lewis Research Center (LeRC) has been developing aerodynamic and aeroelastic analyses for turbomachines. These efforts have resulted in several codes over the years, including applications associated with subsonic, transonic and supersonic regimes, for two- and three- dimensional geometries. A significant collaborative effort between LeRC and Mississippi State University resulted in the development of an unsteady 2-D inviscid flow model for forced response analysis of an isolated cascade called NPHASE (Huff et al., 1991). More recently, Srivastava and Reddy (1994) used NPHASE as the basis for a Multi-Stage Aeroelastic analysis Program 2-Dimensional (MSAP2D) solver for forced response analysis. Although the initial development of NPHASE was considered to be successful, two major shortcomings were identified: (1) treatment of viscous effects, and (2) algorithm efficiency and robustness. NPHASE v2.0 (Loe, 1993) was developed to address these concerns. Srivastava and Reddy then implemented NPHASE v2.0 as the basis for MSAP2D v2.0. But they have not rigorously tested v2.0, including no viscous flow tests. A brief description of NPHASE v1.0 and MSAP2D v1.0 will be followed by a description of NPHASE v2.0 and MSAP2D v2.0. The computational model used to compare with the experimental vane/blade results will be the MSAP2D v2.0 code.

NPHASE and MSAP2D v1.0

The inviscid flow field computations in NPHASE are performed using an implicit finite volume Euler scheme. The discretized integral form of the time-dependent curvilinear Euler equations is obtained by integrating the partial differential equations over a computational volume with the center denoted as (i,j) and changing the resulting volume integral to a surface integral using the divergence theorem

$$\frac{\partial Q}{\partial \tau} + \frac{\delta_i F}{\Delta \xi} + \frac{\delta_j G}{\Delta \eta} = 0 \quad (1)$$

where the central difference operators $\delta_m(\cdot) = (\cdot)_{m+\frac{1}{2}} - (\cdot)_{m-\frac{1}{2}}$, $m=i,j$ imply that the flux vectors are evaluated at the surfaces of a cell. The dissipation aspect of this scheme is improved by a flux-difference split method for the residual based on the solution of approximate Riemann problems with Roe averaging at the cell

faces. To obtain higher order spatial accuracy, a corrective flux is added with a total variational diminishing scheme used to limit the interface flux. The residual term is evaluated with the flux-difference split method and the left hand side operator with the flux vector split scheme, which is approximately factored into the product of two operators. This gives the so-called "2-pass scheme". The geometric conservation law, which prevents spurious source terms due to the motion of the grid is satisfied resulting in a flow solver that is third-order accurate spatially and second-order accurate in time (Huff et al., 1991).

The Euler numerical solution is implemented on an airfoil cascade geometry by a sheared H-grid. A two-dimensional version of the IGB code developed by Beach (1990) for turbomachinery applications is used to generate the H-grid with cells clustered toward the airfoils or coarsened near the inlet and exit boundaries. This grid generator uses a hyperbolic tangent function to determine the spacing upstream and downstream of the cascade and ensures smooth variation in grid spacing.

MSAP2D is an aeroelastic solver for analyzing a multi blade row configuration. Since the current validation is for unsteady vane/blade interaction the structural model will not be described, but the multi blade row boundary condition is discussed. To simulate rotor-stator flow, the solution domain of the two cascade rows is divided into two sets of computational blocks associated with each row, Figure 3. An additional boundary is created between the two cascade rows. The boundary condition utilized between the two cascade computational blocks is an injection method to maintain continuity at this boundary. Due to the relative blade row motion, the grid lines do not always align across the boundary so bi-linear interpolation is used to obtain the proper injection boundary conditions.

NPHASE and MSAP2D v2.0

As stated earlier, the improvements needed for v1.0 of NPHASE are the inclusion of viscous effects and a more efficient and robust algorithm. To address the viscous effects, the thin-layer Navier-Stokes equations have been implemented into the basic NPHASE program such that physical phenomena resulting from viscous effects (including mildly separated flow) could be modeled. The streamwise direction is that associated with the ξ -coordinate and corresponds with distance along the airfoil chord, while the η -coordinate

is defined to be in the blade-to-blade direction. In keeping with the thin-layer approximation, gradients with respect to the ξ -coordinate are neglected. These approximations result in the following governing equation:

$$\frac{\partial Q}{\partial \tau} + \frac{\partial F}{\partial \xi} + \frac{\partial G}{\partial \eta} = \frac{\partial G^d}{\partial \eta} \quad (2)$$

where G^d is the diffusive flux vector containing the derivatives with respect to the η -direction. The algorithm was improved by using an iterative Newton's formulation which was altered to include iterative refinement (also called Gauss-Seidel (GS) iterations) within a given Newton iteration. The effect of increasing the number of GS iterations is to reduce the approximate factorization error. This has been observed to enhance considerably numerical stability. It has been shown (Loe, 1993) that unsteady calculations using v2.0 can be carried out using a time-step eight times greater than that used in v1.0 while retaining the same level of time accuracy. An added advantage is that additional GS iterations are relatively inexpensive (compared to additional Newton iterations) because the residual vector is not recomputed after each GS pass.

Effects of turbulence are modeled through the assumption that the effective viscosity (μ_{eff}) and diffusivity (χ_{eff}) are linear combinations of the molecular and turbulent contributions, or

$$\mu_{eff} = \mu_m + \mu_t \quad , \quad \frac{\chi_{eff}}{Pr} = \frac{\chi_m}{Pr} + \frac{\mu_t}{Pr_t} \quad (3)$$

where $Pr_t = 0.9$ is the turbulent Prandtl number. Two different models are employed, where the model invoked is regionally dependent. For the region bounded by the solid surfaces (i.e., blade-to-blade) the Baldwin-Lomax (1978) model is used. This is a two-layer algebraic eddy viscosity model, where an inner and outer layer eddy viscosity is utilized. The wake region utilizes a model proposed by Thomas (1979), the turbulent viscosity is given by

$$\mu_t = Re \rho l^2 |\omega| \quad (4)$$

where l is the turbulence length scale and $|\omega|$ is the local vorticity.

The computational grid for NPHASE v2.0 is a sheared H-grid with the IGB grid generation code utilized. MSAP2D v2.0 is identical to v1.0 except for the aerodynamic solver. Viscous effects have been included by utilizing NPHASE v2.0, which models the thin-layer Navier-Stokes equations. Also, the efficiency and robustness of MSAP2D v2.0 is improved by using NPHASE v2.0. The procedures for

incorporating the multi blade analysis is the same as v1.0. The reader is referred to the earlier discussion of this method.

Preliminary Computational Results

Before MSAP2D could be used to study the unsteady aerodynamics of the SMI's vane/blade interaction, the code had to be ported to the workstation environment at CARL. MSAP2D arrived from LeRC coded to run on a Cray supercomputer utilizing Cray specific logic. Once the conversion and verification was complete, MSAP2D was ready to be used for CARL's SMI configuration. Because MSAP2D is a 2-D code the 50% spanwise location was selected for initial analysis. Figure 4 shows the IGV, rotor, and stator blade geometries for the 50% spanwise location of the SMI rig.

With the experimental geometry defined, sheared H computational grids were developed for both the IGV and rotor blade geometry using the IGB grid generation code. Since NPHASE is a subset of the MSAP2D multi blade row aeroelastic code, NPHASE will be used to analyze the steady viscous, convergence and grid independence studies. The analysis shall proceed as follows. The IGV geometry is studied for steady flow conditions to determine the computational grid required to resolve the viscous effects. Once grid independence is achieved then the cascade pressure loss coefficient will be calculated for several different mass flow rates and compared with design calculations supplied by Pratt and Whitney (Cheatham, 1991) for the IGV blades. Similarly, the rotor geometry will be analyzed. The computational analysis will then move to the MSAP2D code to incorporate the unsteady vane/blade interaction effects. The MSAP2D will need to account for the proper number of IGV blades and rotor blades. As stated earlier, the experimental tests are planned for the IGV configuration with 24 blades. The rotor has 33 blades, so to exactly model the IGV/rotor section with the time accurate MSAP2D code would require solving the Navier-Stokes equations for all 57 blades. Therefore, it is necessary to reduce these numbers while still properly modeling the system. With the use of periodic boundary conditions, the number of wake generators can be cut to 2 with the number of rotor blades reduced to 3. This configuration is accurate for a 36 bladed rotor, therefore it is necessary to reduce the size of the rotor blades and the solidity by a factor 33/36 in order

to accurately model the test geometry. Figure 5 shows the general 2/3 IGV/rotor geometry and a possible computational grid.

Conclusion

The preliminary research has been completed for an experimental and computational analysis of the influence of a transonic compressor rotor on the upstream IGV blades wake characteristics. The primary experimental data will be collected by using 20 miniature high response pressure transducers mounted on the surface of the IGV blades. Two spanwise locations will be studied to analyze possible three dimensional effects of the unsteady vane/blade interaction. The experimental investigation will be complimented by a detailed computational analysis. The computational analysis will utilize a time-accurate nonlinear unsteady thin-layer Navier-Stokes code. The computational code is designed to model unsteady vane/blade interaction. Both the IGV blades and the rotor blades will be modeled and allowed to move relative to each other like in the experiment. The current time schedule is for this project to be completed by June, 1997.

Acknowledgments

First, the authors would like to thank all the CARL personnel for all the help and advice given during this study. We could not have asked for a better research environment and are anticipating continued interaction over the next 10 months to obtain the final results of this study. We also acknowledge T. S. R. Reddy and T. A. Beach from NASA Lewis Research Center for providing the computational codes. This work was sponsored by the AFSOR SFRP.

References

- Baldwin, B., and Lomax, H., 1978, "Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows," AIAA paper no. 78-257.
- Beach, Timothy A., 1990, "An Interactive Grid Generation Procedure for Axial and Radial Flow Turbomachinery," NASA CR-185167.

Cheatham, J.G., 1991, "Stage Matching Investigation," Final Report-Contract No. F33615-88-C-2801, United States Air Force, Wright Research & Development Center, Wright-Patterson AFB.

Dring, R.P., Joslyn, H.D., Hardin, L.W., and Wagner, J.H., 1982, "Turbine Rotor-Stator Interaction," Journal of Engineering for Power, Vol. 104, pp. 729-742.

Dunn, M.G., Bennett, W.A., Delaney, R.A., and Rao, K.V., 1992, "Investigation of Unsteady Flow Through a Transonic Turbine Stage: Data/Prediction Comparison for Time-Averaged and Phase-Resolved Pressure Data," Journal of Turbomachinery, Vol. 114, pp. 91-99.

Giles, M.B., 1988, "Stator/Rotor Interaction a Transonic Turbine," AIAA Paper 88-3093.

Huff, Dennis L., Swafford, Timothy W., and Reddy, T.S.R., 1991, "Euler Flow Predictions for an Oscillating Cascade Using a High Resolution Wave-Split Scheme," NASA TM-104377.

Loe, David Henderson, 1993, "Numerically Stable Unsteady Viscous Flow Predictions for Oscillating Cascades," Masters Thesis, Mississippi State University, Mississippi.

Rai, Man Mohan, 1989, "Three-Dimensional Navier-Stokes Simulations of Turbine Rotor-Stator Interaction; Part I-Methodology," Journal of Propulsion, Vol. 5, No. 3, pp.305-311.

Rao, K., and Delaney, R., 1990, "Investigation of Unsteady Flow Through Transonic Turbine Stage; Part 1: Analysis," AIAA paper no. 90-2408.

Rao, K.V., and Delaney, R.A., 1992, "Vane-Blade Interaction in a Transonic Turbine Part 1- Aerodynamics," AIAA paper no. 92-3323.

Srivastava, R., and Reddy, T.S.R., 1994, "Forced Response Analysis Using a Multistage Euler Aeroelastic Solver," AIAA paper no. 94-0739.

Suddhoo, A., Giles, M.B., and Stow, P., 1991, "Simulation of Inviscid Blade Row Interaction Using a Linear and a Non-Linear Method," AIAA paper no. 91-7049.

Thomas, P.D., 1979, "Numerical Method for Predicting Flow Characteristics and Performance of Nonaxisymmetric Nozzles -- Theory", NASA CR-3147, Langley Research Center.

Verdon, Joseph M., 1992, "Unsteady Aerodynamic Methods for Turbomachinery Aeroelastic and Aeroacoustic Applications," AIAA paper no. 92-0011.

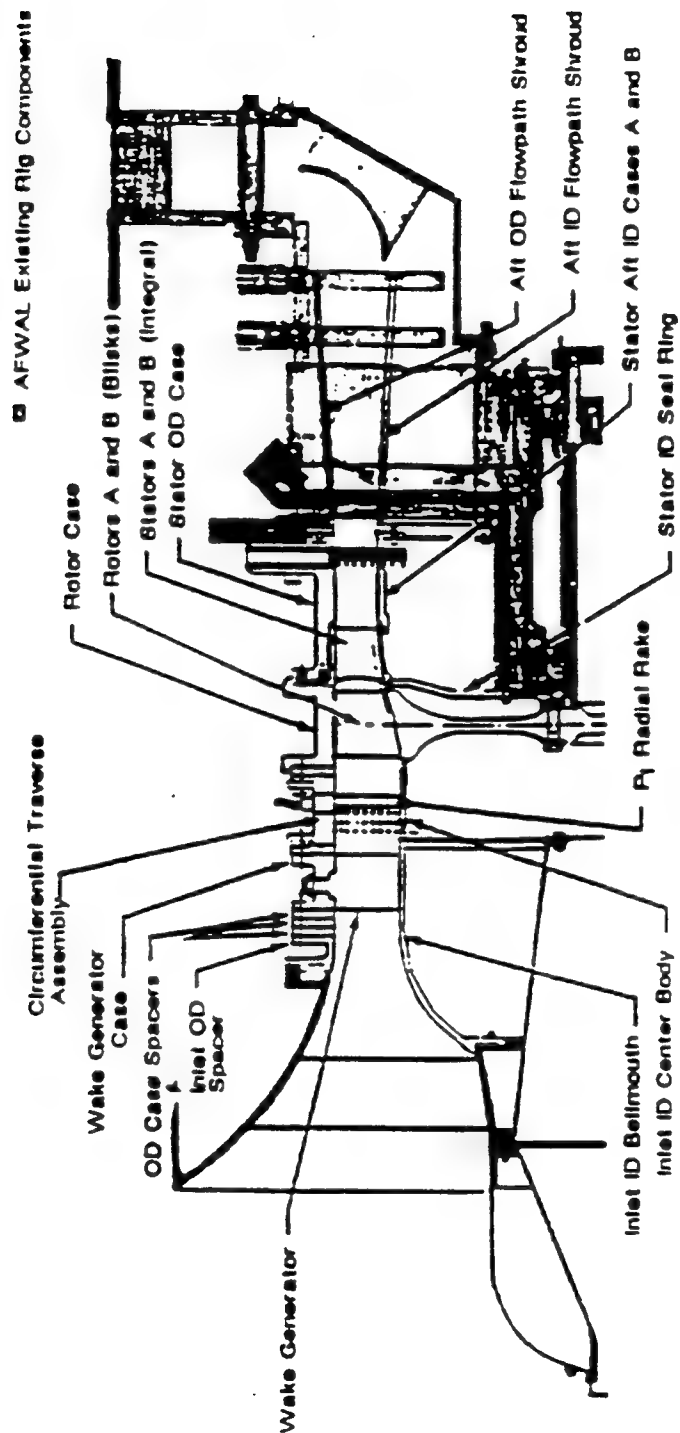
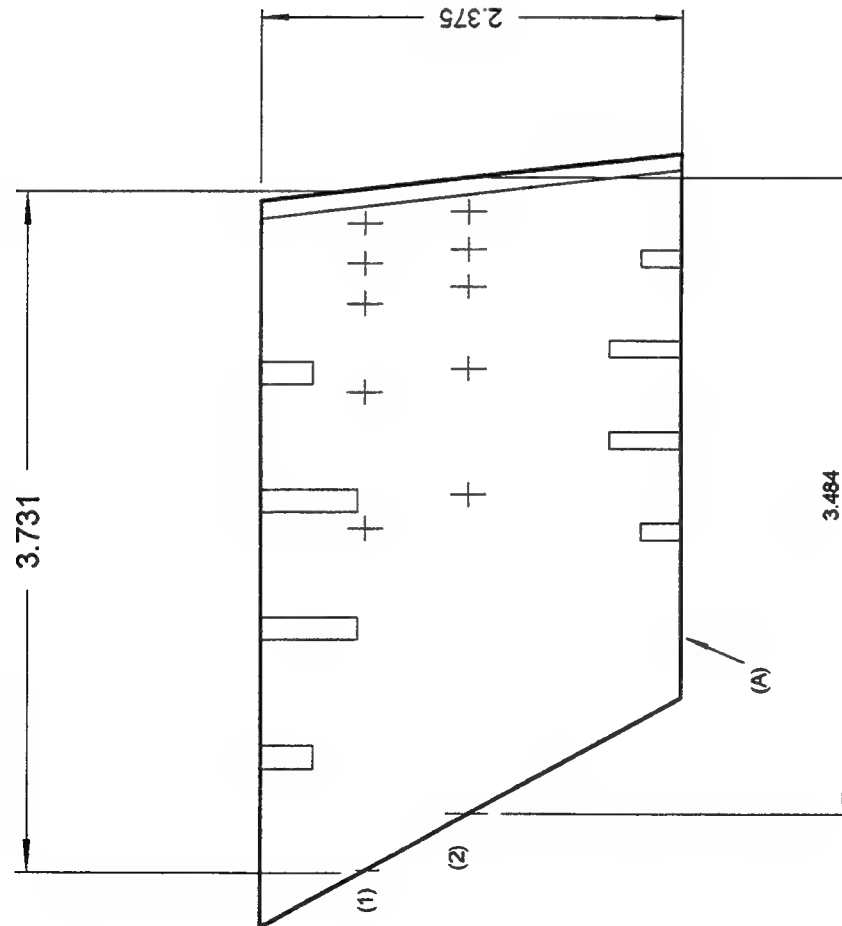


Figure 1: Schematic of the CARL SMI rig (note: the term wake generator is the same as IGV).



Chordwise Locations of Transducers

UPPER - - from ref. point (1)

95%(3.544 in.)
89%(3.321 in.)
83%(3.097 in.)
70%(2.612 in.)
50%(1.866 in.)

LOWER - - from ref. point (2)

95%(3.294 in.)
89%(3.086 in.)
83%(2.877 in.)
70%(2.427 in.)
50%(1.734 in.)

Spanwise Locations - - from ref. line (A)

50%(1.188 in.)
75%(1.781 in.)

Figure 2: Drawing of IGV surface pressure transducer locations.

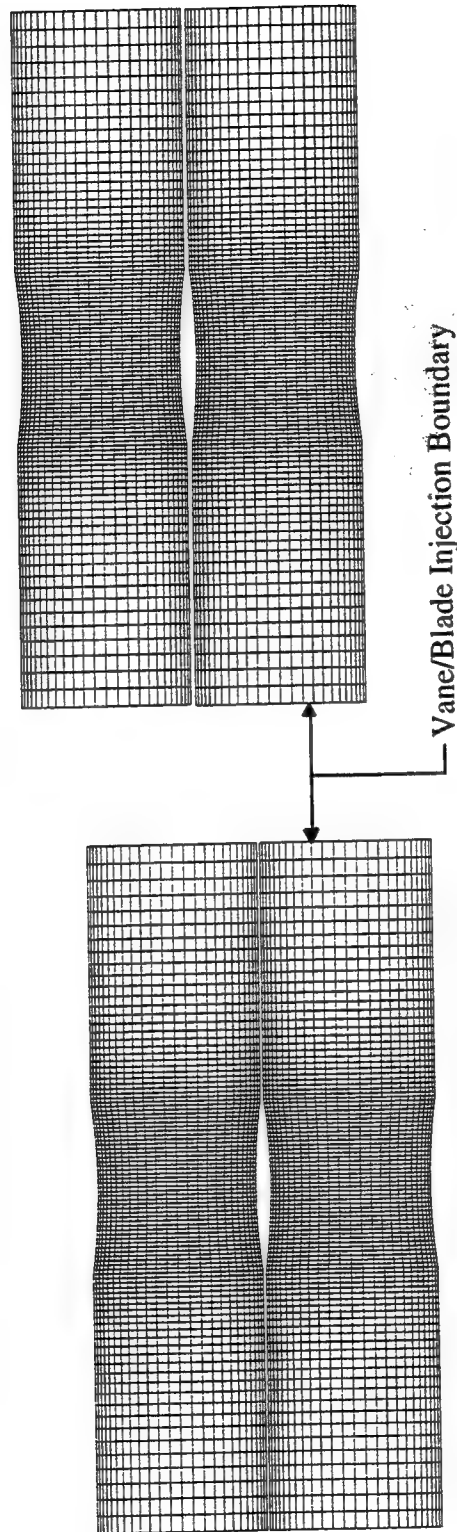


Figure 3: Injection computational boundary for vane/blade interaction analysis.

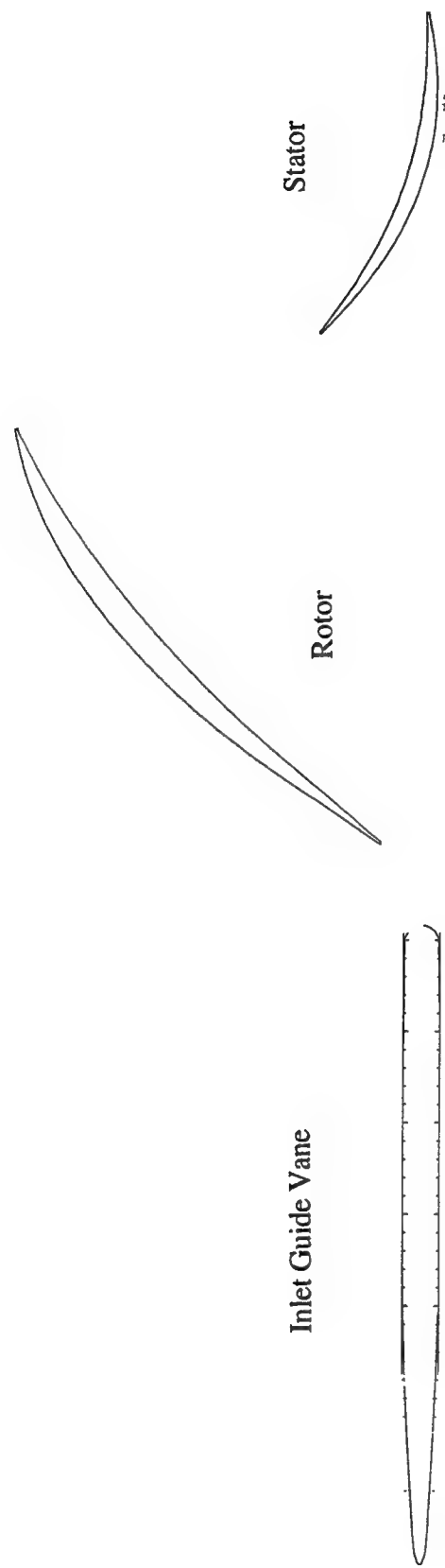


Figure 4: Two dimensional mid-span SMI core rotor flow path geometry.

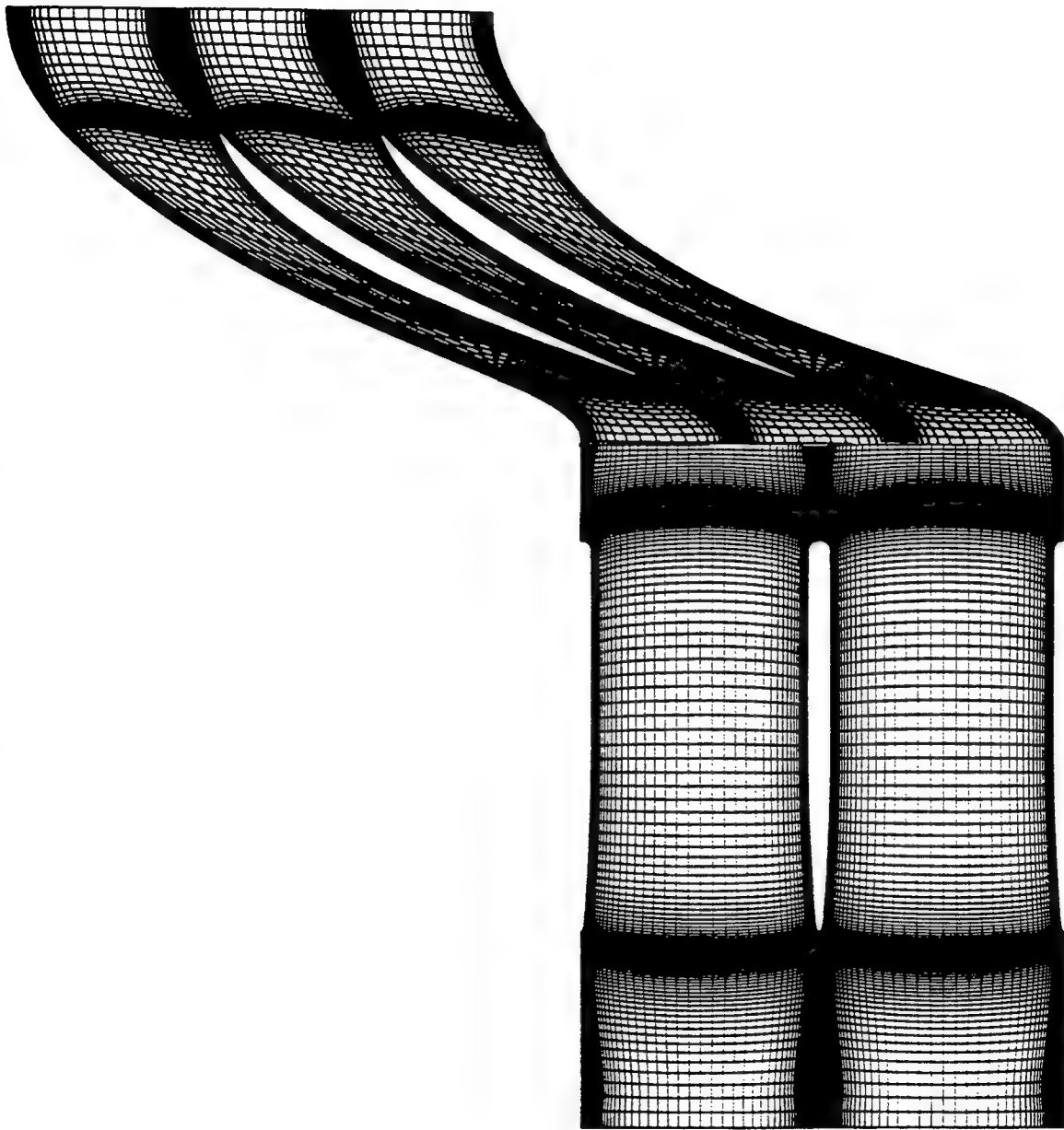


Figure 5: Example computational grid for use with MSAP2D in analyzing the SMI core compressor.

INTEGRATED AEROELASTIC SMART STRUCTURE CONTROL USING ACCELERATION MEASUREMENT FEEDBACK

R. K. Yedavalli

Professor

Department of Aerospace Engineering,

Applied Mechanics, and Aviation.

The Ohio State University

Columbus, Ohio 43210-1276

Final Report for:

Summer Faculty Research Program

Wright Laboratory; Wright-Patterson AFB, Ohio 45433

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, DC

and

Wright Laboratory

September 1996

INTEGRATED AEROELASTIC SMART STRUCTURE CONTROL USING ACCELERATION MEASUREMENT FEEDBACK

R. K. Yedavalli

Professor

Department of Aerospace Engineering,
Applied Mechanics, and Aviation.

The Ohio State University

Columbus, Ohio 43210-1276

Abstract

In the research during the Summer Faculty Research Program, the aspect of designing control systems for controlling a composite plate structure subject to aeroelastic loading using piezoelectric actuators and sensors and acceleration measurements was studied and new controller designs are developed. The nature of the sensing variables and the distributed nature of smart material actuation and sensing is exploited and accommodated in the control design algorithms. The controllers designed include a state estimate based feedback as well as a pure measurement feedback using various measurements related to motion variables such as accelerations, velocities and displacements. The proposed methods provide different norms for estimator and controller gains, thereby allowing more flexibility in gain magnitudes and selection of sensors. Then the controller gains and the structural weight are optimized by using integrated structure/control optimization schemes. Finally, preliminary designs are initiated for active flight control with smart deformable wing structures.

INTEGRATED AEROELASTIC SMART STRUCTURE CONTROL USING ACCELERATION MEASUREMENT FEEDBACK

R. K. Yedavalli

1 Introduction:

In recent years, design and control of smart structures has become an active topic of research due to the significant potential of these structures to actively deform for control purposes thereby improving the performance of various types of flight vehicles. In particular, piezoelectric actuation and sensing in the form of Micro Electro Mechanical Systems (MEMS) is gaining lot of attention due to the distributed nature of these 'smart' materials. Typically, the modeling of the vibrational motion of these structures is done using finite element procedures, which yield, in the linear range, simultaneous, ordinary differential equations of second order in the 'configuration or generalized' coordinates, which we call the Matrix Second Order Systems involving Mass, Damping and Stiffness matrices. Traditionally, analysis and design of control systems for these structures is done by transforming the above system to a state space form. However, in State Space coordinates, physical insight of mass, damping and stiffness terms is lost and also acceleration feedback cannot be implemented directly. Acceleration feedback, directly using the accelerometers, is particularly desirable and attractive in control design because accelerometers are frequently used in structural control (for example, many commercially available MEMS are accelerometers). Also, the 'distributed' nature of piezoelectric 'smart' actuation and sensing is especially attractive in control design because it allows the control designer additional degrees of sensing and actuation. In many of present control design techniques, no effort is made to exploit the special nature of the distributed nature of these smart structures as well as the availability of acceleration information. The main viewpoint taken in this research is that the task of extracting improved performance from the existing smart materials lies not entirely on developing improved materials but rather on the innovative and creative efforts in the areas of modeling, design and control algorithms. So in this research, new efforts are undertaken to design controllers that explicitly take the nature

of sensing (such as acceleration measurements) into account and also provide more flexibility in the use of various combination of sensors thereby providing improved alternatives to State Space design methods. In addition, the proposed methods provide different norms for the gains, thereby allowing trade off between control gain and sensor usage. The controller design methods are essentially based on eigenvalue placement. Another relevant topic of research that was addressed in this research is the integrated structure/control optimization in the presence of unmodeled dynamics and real parameter variations. The nominal control designs proposed in the previous discussion are now optimized for weight and closed loop eigenvalues taking into account the inevitable presence of uncertainties such as 'unmodeled dynamics' (i.e. unstructured uncertainty) and 'real parameter variations' (i.e. structured uncertainty). The proposed algorithm for this purpose is efficient and simple in the sense that it uses the existing computational and optimization algorithms without any additional complexity. Encouraged by the success of these methodologies, finally a preliminary research effort was undertaken to design control logic for the static aeroelastic roll maneuvering control of an aircraft using active deformations of smart wing structures.

2 Control Design for Smart Structural Systems:

The typical model of a discretized structure in linear range is described by the Matrix Second Order System of the form,

$$M\ddot{u} + E\dot{u} + Ku + Df = 0$$

where M , E and K are the Mass, Damping and Stiffness matrices of dimension $n \times n$ and D is the control distribution matrix of dimension $m \times m$ and f is the m vector of control variables and u is the n vector of generalized physical coordinates. Typically, control design in matrix second order form is carried out using pole (or eigenvalue) placement techniques. However, the available methods [1], [2] of eigenvalue assignment either have limitations or need time consuming computations (eg. singular value decomposition). methods of designing control gains directly in matrix second order design space are mostly restricted to symmetric, positive definite mass, damping and stiffness matrices. In this paper, no such restriction of symmetry is assumed on system matrices (M , E & K). Since designing control gains to achieve desired closed loop eigenvalue placement directly using matrix second order design framework for general, nonsymmetric, indefinite mass, damping,

stiffness matrices is a formidable task at the present stage of research, at this point , a compromise is suggested by first using state-space methods as a baseline design of controllers and estimators (which involve information of velocities and displacements only) and then using these baseline designs, different controllers and estimators are obtained which now provide acceleration feedback directly. In fact the proposed method even allows freedom in selecting different types of sensors and controllers. If singularity does not occur in mass,damping, and stiffness matrices , according to the proposed method, feedback of any two kinds of motion variables can accomplish eigenvalue assignment.

2.1 Modeling of a Composite Plate Subject to Aerodynamic Loads:

The dynamic analysis of transverse vibration of a piezoelectric plate subjected to aerodynamic load is reviewed here. The coupled equation of motion for two dimensional cantilever composite plate with distributed piezoelectric PVDF layer on one side is derived using variational method [3]. The plate is subjected to a uniform airstream of velocity V in the direction of the positive y -axis (Please see fig 1).

The approximate out-of-plane deflection w of the plate and the electric potential Φ in generalized coordinates, are expanded as

$$w(x, y, t) = \sum_{i=1}^n N_i(x, y) u_i(t) \quad (1)$$

$$\Phi(x, y, z) = \sum_{i=1}^p \phi_i(x, y, z) u_{ei} \quad (2)$$

where $u(t)$ is the generalized displacement coordinate and u_e are generalized electrical potential and n is the dimension of $u(t)$ and u_e . The mechanical shape functions N_i satisfy the geometric boundary condition while the electric potential shape functions ϕ_i are consistent with prescribed voltage/charge boundary conditions. A four-term approximate deflection expression [4] is suggested to determine the first and second spanwise bending modes, the first and second torsion modes of the rigid rotation of the plate mean chord about the midchord. One shape function for electric potential is used in order to obtain a actuator matrix. The plate is the same as the laminated composite lifting surfaces in [5] and 0.2mm PVDF material is distributed on its upper surface. The mass matrix M_s and stiffness matrix K_s of composite plate are obtained from [5]

while the physical properties of piezoelectric layer can be found from [6]. Applying variational method we get mass matrix M_{pz} and stiffness matrix K_{pz} for this electroelastic system

$$M_{pz} = \int_{t_b}^{t+t_b} \int_{-0.5c}^{0.5c} \int_0^l \rho_{pz} N^T N dx dy dz \quad (3)$$

$$K_{pz} = \int_{t_b}^{t+t_b} \int_{-0.5c}^{0.5c} \int_0^l b^T D_{pz} b dx dy dz \quad (4)$$

where

$$b^T = \left[-\frac{\partial^2 N}{\partial x^2} \quad -\frac{\partial^2 N}{\partial y^2} \quad 2\frac{\partial^2 N}{\partial x \partial y} \right] \quad (5)$$

l is the length of the plate; c is the width of the plate; t_b is the half thickness of the plate; t is the thickness of applied PVDF material. The coupling matrix $K_{u\phi}$ and capacitance matrix $K_{\phi\phi}$ as follows:

$$K_{u\phi} = \int_{t_b}^{t+t_b} \int_{-0.5c}^{0.5c} \int_0^l b^T e^T b_\phi dx dy dz \quad (6)$$

$$K_{\phi\phi} = \int_{t_b}^{t+t_b} \int_{-0.5c}^{0.5c} \int_0^l b_\phi^T \varepsilon b_\phi dx dy dz \quad (7)$$

where

$$b_\phi = \frac{\partial \phi}{\partial z}$$

Assuming we use i separated electrodes with charge input at the points (x_i, y_i) , here, $i = 1$,

$$q(x, y, t) = q(t) \delta(x - x_i) \delta(y - y_i) \quad (8)$$

The electric control force Q is defined as:

$$Q(t) = \int_{t_b}^{t+t_b} \int_{-0.5c}^{0.5c} \int_0^l \phi(x, y, z) q(x, y, t) dx dy dz = \Gamma q(t) \quad (9)$$

where $\Gamma = \phi_i(x_i, y_i)$

In this investigation, one accelerometer is installed at $(\frac{c}{2}, \frac{3l}{4})$ as sensor, therefore, the sensor matrix C is as follows

$$C = \left[N_1(\frac{c}{2}, \frac{3l}{4}) \quad N_2(\frac{c}{2}, \frac{3l}{4}) \quad N_3(\frac{c}{2}, \frac{3l}{4}) \quad N_4(\frac{c}{2}, \frac{3l}{4}) \right] \quad (10)$$

where N_i are structural shape functions described as before.

2.2 Aerodynamic loads:

Following [3], [7] and using the first order high Mach number approximation to the linear potential flow theory, the work done by the external aerodynamic forces is given by

$$W_a = - \int_A \frac{2Q_a}{V\beta} \left(V \frac{\partial w}{\partial x} + \left(\frac{M^2 - 2}{M^2 - 1} \right) \frac{\partial w}{\partial t} \right) w dA \quad (11)$$

where $\beta = \sqrt{M^2 - 1}$ and M is Mach number, $Q_a = \frac{\rho V^2}{2}$ is the dynamic pressure. Applying variational method, we obtain the aerodynamic damping matrix A_1 and aerodynamic stiffness matrix A_2 . It may be noted that A_2 is a nonsymmetric matrix.

$$A_1 = \frac{2Q_a(M^2 - 2)}{V(M^2 - 1)^{1.5}} \int_{-0.5c}^{0.5c} \int_0^l N^T N dx dy \quad (12)$$

$$A_2 = \frac{2Q_a}{(M^2 - 1)^{0.5}} \int_{-0.5c}^{0.5c} \int_0^l N^T N_x dx dy \quad (13)$$

Finally, we obtain the equations of motion for a piezoelectrically coupled electromechanical composite plate in a flow field as follows:

$$\begin{bmatrix} M_s + M_{pz} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \ddot{u}_e \end{bmatrix} + \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{u}_e \end{bmatrix} + \begin{bmatrix} K_s + K_{pz} + A_2 & -K_{u\phi} \\ K_{\phi u} & K_{\phi\phi} \end{bmatrix} \begin{bmatrix} u \\ u_e \end{bmatrix} = \begin{bmatrix} 0 \\ \Gamma q \end{bmatrix} \quad (14)$$

Reducing electric potential from the system equation, we get decoupled displacement and electric potential equations as:

$$(M_s + M_{pz})\ddot{u} + A_1\dot{u} + (K_s + K_{pz} + A_2)u - K_{u\phi}u_e = 0 \quad (15)$$

$$K_{\phi u}u + K_{\phi\phi}u_e = \Gamma q \quad (16)$$

From this, we have

$$u_e(t) = K_{\phi\phi}^{-1}\Gamma q - K_{\phi\phi}^{-1}K_{\phi u}u(t) \quad (17)$$

Substituting eq.(18) into eq.(16), we have

$$(M_s + M_{pz})\ddot{u} + A_1\dot{u} + (K_s + K_{pz} + A_2 + K_{u\phi}K_{\phi\phi}^{-1}K_{\phi u})u = K_{u\phi}K_{\phi\phi}^{-1}\Gamma q \quad (18)$$

Rewrite the above as $M\ddot{u} + E\dot{u} + Ku + Df = 0$

where $M = M_s + M_{pz}$, $E = A_1$, $K = K_s + K_{pz} + A_2 + K_{u\phi}K_{\phi\phi}^{-1}K_{\phi u}$, $D = -K_{u\phi}K_{\phi\phi}^{-1}\Gamma$ and $f = q$.

It may be noted that since A_2 is not symmetric, K is not symmetric either. Note that the above matrix second order system equation is quite different from the typical equation generated by pure structural dynamics, in which case, we normally have symmetric, positive definite mass and stiffness matrices and no damping matrix. Here the damping matrix A_1 is generated due to the aeroelastic damping and the overall stiffness matrix K includes the aeroelastic stiffness matrix A_2 which could make K nonsymmetric and indefinite. Thus the development of controller to achieve desired performance is not routine and straightforward.

3 Controller and Estimator Designs for Matrix Second Order Systems:

In this section, we propose a control design methodology for the above general(nonpositive definite, nonsymmetric) matrix second order systems. Since acceleration measurements are typically available (as in MEMS), we strive to propose a control logic that explicitly provides acceleration feedback. So we start with the equation of motion of physical systems in matrix second order form given by,

$$M\ddot{u} + E\dot{u} + Ku + Df = 0$$

and propose a generalized force f as given by:

$$f = G_a\ddot{u} + G_v\dot{u} + G_d u$$

Since in reality not all the physical variables are sensed, we assume a measurement equation of the form

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} C & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & C \end{bmatrix} \begin{bmatrix} u \\ \dot{u} \\ \ddot{u} \end{bmatrix} \quad (19)$$

where the dimension of the measurement vector Z is $3l$ and the given sensor matrices C , are $l \times n$ matrices which represent the locations of accelerometers. Once the accelerations are sensed, using integrators one can obtain the corresponding velocities and displacements. This philosophy

of using accelerometers explicitly for feedback purposes and using integrations to obtain other lower derivative motion variables is the main feature of our proposed methodology and is in direct contrast to the state space design methodology wherein only displacement and/or velocity measurements are assumed and feedback of acceleration information is accomplished as a differentiation of the velocity information (that essentially amounts to a Proportional plus Derivative Controller) which we believe not only suffers from noise problems but also is uneconomical, inefficient and unproductive.

3.1 Estimator/ Controller Designs:

We now present a method to design an Estimator based on the above measurements and then a controller that uses these estimated variables for feedback. We suppose that acceleration sensors are installed on structural system and velocity and/or displacement are obtained by integrating acceleration information. We first propose a baseline estimator (using only displacement and velocity sensors) for matrix second order systems of the following form:

$$M\ddot{u} + (E - L_v C)\dot{u} + (K - L_d C)u + L_v z_2 + L_d z_1 + Df = 0 \quad (20)$$

Letting $e = u - \hat{u}$ as measurement error, we have

$$M\ddot{e} + (E - L_v C)\dot{e} + (K - L_d C)e = 0 \quad (21)$$

The task for estimator design is first to find baseline estimator gains L_v and L_d . Using the fact the a matrix and its transpose have the same eigenvalues, we create a new system, given by

$$M^T \ddot{e} + (E - L_v C)^T \dot{e} + (K - L_d C)^T e = 0 \quad (22)$$

and putting the above in state space form as

$$\dot{q} = \bar{A}q + \bar{B}\bar{L}q \quad (23)$$

we design the baseline estimator gains $\bar{L} = [L_d^T \ L_v^T]$ such that matrix $(\bar{A} + \bar{B}\bar{L})$ is stable by using any pole placement methods or optimization methods as long as the pair of \bar{A} and \bar{B} is controllable. Once we obtain a satisfactory baseline design, the following estimator design is carried out that

provides the same performance as the above but offers flexibility in sensor selection and estimator gain magnitudes.

When K is nonsingular, letting $y^T = [C\ddot{u} \ C\dot{u}] = [y_a \ y_v]$ (i.e. no displacement is measured), the estimator can be expressed as

$$(M - L_{a1})\ddot{\hat{u}} + (E - L_{v1}C)\dot{\hat{u}} + K\hat{u} + L_{a1}z_3 + L_{v1}z_2 + Df = 0 \quad (24)$$

where

$$L_{a1}^T = -L_d^T(K^T - C^T L_d^T)^{-1}M^T \quad (25)$$

and

$$L_{v1}^T = L_v^T - L_d^T(K^T - C^T L_d^T)^{-1}(E^T - C^T L_v^T) \quad (26)$$

Similarly, when E is nonsingular, letting $y^T = [C\ddot{u} \ Cu] = [y_a \ y_d]$ (i.e. no velocity is measured), the estimator can be expressed as

$$(M - L_{a2})\ddot{\hat{u}} + E\dot{\hat{u}} + (K - L_{d2})\hat{u} + L_{a2}z_3 + L_{d2}z_1 + Df = 0 \quad (27)$$

where

$$L_{a2}^T = -L_v^T(E^T - C^T L_v^T)^{-1}M^T \quad (28)$$

and

$$L_{d2}^T = L_d^T - C^T L_v^T(E^T - C^T L_v^T)^{-1}(K^T - C^T L_d^T) \quad (29)$$

Similarly, expressions for gains of other combinations can be obtained. Also similar logic is used to design controller gains. Once baseline feedback gains of velocity and displacement are available, we can use above equations to obtain corresponding acceleration/velocity and acceleration/displacement feedback gains. This can provide more flexibility in implementation of sensors and controllers. Then it turns out that from the proposed structures for the controller and estimator as described above, the desirable Separation Principle holds for the present matrix second order systems also. The eventual control law for the matrix second order system is given by:

$$M\ddot{u} + E\dot{u} + Ku + Df = 0(Plant) \quad (30)$$

$$f = G_a\ddot{\hat{u}} + G_v\dot{\hat{u}} + G_d\hat{u}(Control) \quad (31)$$

where \hat{u} is estimated displacement.

In the next section, we present a technique for simultaneous structure-control optimization with robustness against unmodeled dynamics and real parameter variations.

4 Integrated Structure-Control Optimization with Robustness to Unmodeled Dynamics and Real Parameter Uncertainty:

4.1 Introduction and Motivation:

It is well known that simultaneous optimization of Structure, Control and Aerodynamics is important for improved designs in flight vehicles. Obviously one needs Multidisciplinary Design and Analysis Methods to achieve this goal. In fact this topic has been a very active research area for the last decade or more. In this area, significant contributions were made by WL researchers [8], [9] In this research, uncertainty characterization is critical for control design purposes. There are essentially two ways in which the uncertainty can be captured in the problem formulation. One is the uncertainty due to 'Unmodeled Dynamics' which we label as 'Unstructured Uncertainty'. Control design methods which protect against unmodeled dynamics error are the now popular H-Infinity Design, and the more recent Mixed H-two/ H-Infinity Design methodologies. A useful contribution in this area is [10]. As far as the structured real parameter uncertainty is concerned, the research is focussed more on analysis techniques than on design. However in a realistic application, we encounter both of these uncertainties and we need a method that accommodates this combined uncertainty. Towards this direction in this section, a method developed during the Summer Research program is presented. The controller is a measurement feedback robust controller that has protection against both unmodeled dynamics as well as structured real parameter uncertainty.

4.2 Integrated Structure Control Optimization: Mixed H_2/H_∞ Measurement Feedback Control Design

Consider the following state space representation of the structural system given by

$$\text{System: } \dot{x} = Ax + Bu + Ew$$

$$\text{Measurements: } z = Hx$$

$$\text{Output: } y = Cx$$

$$\text{Control: } u = Fz$$

where

A is $n \times n$, B is $n \times m$, D_1 is $n \times p$, K is $m \times l$, C is $l \times n$, R_1 is $n \times n$, R_2 is $m \times m$.

and w : external disturbance vector, for white noise, and u : control vector

$$\begin{aligned} PI &= \lim_{n \rightarrow \infty} \frac{1}{t} [y^T \bar{Q} y + u^T R_2 u] = \lim_{n \rightarrow \infty} \frac{1}{t} [x^T C^T \bar{Q} C x + u^T R_2 u] \\ &= \lim_{n \rightarrow \infty} \frac{1}{t} [x^T R_1 x + u^T R_2 u] \end{aligned}$$

Let $y_\infty = C_\infty x$ be the outputs for which we need H-infinity norm attenuation. Now let

$$u = Fz = FHx$$

The closed loop system is given by

$$\dot{x} = (A + BFH)x + Ew = A_{CL}x + Ew$$

$$y = Cx$$

$$y_\infty = C_\infty x$$

and the performance index is given by

$$J_L = \lim_{n \rightarrow \infty} \frac{1}{t} [x^T (R_1 + H^T F^T R_2 F H) x] = \text{Trace} P_2 (R_1 + H^T F^T R_2 F H)$$

where P_2 is obtained from

$$(A + BFH)P_2 + P_2(A + BFH)^T + EE^T = 0$$

The transfer matrix between the output y_∞ and the disturbance w is given by

$$T_{y_\infty w} = C_\infty (SI - A_{CL})^{-1} E$$

and we require for a positive scalar γ that

$$\|T_{y_\infty w}\|_\infty < \gamma$$

The mixed H_2/H_∞ formulation involves minimizing an index that is an upperbound to J_L , i.e.

$$\tilde{J} = \text{tr}[P_\infty(R_1 + H^T F^T R_2 F H)] \geq J_L$$

where

$$(A + BFH)P_\infty + P_\infty(A + BFH)^T + \gamma^{-2}P_\infty(R_{1\infty}^-)P_\infty + EE^T = 0$$

Let us now add Real Parameter Variations to the system under consideration:

$$\dot{x} = (A + \Delta A)x + (B + \Delta B)u + Ew$$

$$z = Hx$$

$$u = Fz = FHx$$

Let

$$\Delta A : |\Delta A_{ij}| \leq a_{ij\max} \Rightarrow A_m = [a_{ij}]_{\max}$$

$$\Delta B : |\Delta B_{ij}| \leq b_{ij\max} \Rightarrow B_m = [b_{ij}]_{\max}$$

Then it is known that the 'degree of stability, α ' serves as protection against real parameter variations [11]. In other words, if

$$\alpha = - \min_i \text{Re} \lambda_i(A + BFH) > \|(M^{-1})^+[A_m + B_m(FH)^+]M^+\|_2$$

then the closed loop system is guaranteed to be stable in the presence of these parameter variations. For simplicity, we assume $\Delta B = 0$ and obtain the needed necessary conditions for optimal control as

$$(A + BFH + \alpha I)P_\infty + P_\infty(A + BFH + \alpha I)^T + \gamma^{-2}[P_\infty R_{1\infty}^- P_\infty + EE^T] = 0 \quad (32)$$

$$(A + BFH + \alpha I + \gamma^{-2}P_\infty R_{1\infty}^-)^T X + X(A + BFH + \alpha I + \gamma^{-2}P_\infty R_{1\infty}^-)$$

$$+ R_1 + H^T F^T R_2 F H = 0 \quad (33)$$

$$F = -R_2^{-1} B^T (X P_\infty + P_\alpha Y) H^T (H P_\infty H^T)^{-1} \quad (34)$$

$$(A + BFH + \alpha I)^T Y + Y(A + BFH + \alpha I) = 0 \quad (35)$$

$$(A + BFH + \alpha I)^T P_\alpha + P_\alpha (A + BFH + \alpha I) + 2Q = 0 \quad (36)$$

The optimization involves solving these above equations simultaneously with respect to the design variables and because of space limitations, details are not provided here.

5 Roll Maneuvering in Flexible Wings with Smart Material Excitation:

In order to enhance Flight Vehicle Performance (such as Maneuverability, Drag and Radar Signature) distributed actuation via smart materials is proposed to be a feasible technique. This type of distributed actuation eliminates the need for discrete control surfaces thereby improving radar signature. The strategy is to induce active deformation of the structures through smart material excitation which provides significant control on the aerodynamic forces and moments. To this extent some preliminary research was carried out in [12]. In this section, we attempt to gain better insight to this problem by first carrying out the analysis using simple analytical models. Thus the intent of the analysis is to understand the dynamics of 'roll maneuvers' and obtain the relationship between the desired 'roll rates' and the corresponding excitation forces and moments to be generated by the smart material embedded in the structure. We provide a preliminary study undertaken to study this problem in this section.

5.0.1 Simple Static Aeroelastic Problem:

The following assumptions are made (please see figure 2): A1: Slender Straight Wings with elastic axis perpendicular to Fuselage Center Line. A2: No Camber bending, only Pure Torsional Twist.

Note that we propose to generate the entire roll through pure elastic twist (with no discrete control surfaces). Thus the implication is that an arbitrary twist (torsional displacement) on both sides of the wing always results in a rolling moment with some roll rate p . In other words, if the

twist is entirely symmetrical there would be no roll and whatever roll is generated is due to the antisymmetrical components of the wing twist and the associated aerodynamic lift distribution. Note that the general arbitrary twist is a superposition of symmetrical and antisymmetrical components of the twist. This also means that there is no roll whenever the wings are straight and level.

Thus, the local angle of attack under a roll maneuver can be written as

$$\alpha(y) = \alpha^r(y) + \theta(y) - \frac{py}{U} \quad (37)$$

so that

$$\alpha_{symm} = \alpha^r(y) + \theta_{symm} \quad (38)$$

$$\alpha_{antisymm} = \theta_{antisymm} - \frac{py}{U} \quad (39)$$

where $\frac{py}{U}$ is the induced angle of attack due to roll rate p and $\alpha^r(y)$ is the local angle of attack measured from zero lift, not including elastic twist and spanwise aerodynamic induction effects. It may contain such terms as wing attitude, geometric twist, aerodynamic twist resulting from deflected control surfaces or induced angle of attack due to a gust. Note that $\theta(y)$ is the pure elastic twist.

Since, in our case, we do not intend to use any external control surfaces such as ailerons, as well as to focus our attention on the problem of relating desired roll rate with pure elastic twist, from now on we take $\alpha^r(y) = 0$. Thus we have

$$\alpha(y) = \theta(y) - \frac{py}{U} \quad (40)$$

Accordingly, the local lift coefficient distribution can be written as

$$C_L(y) = C_L^e(y) + C_L^{ir}(y) \quad (41)$$

where

$C_L^e(y)$ = Local Lift Coefficient resulting from the elastic twist $\theta(y)$

$C_L^{ir}(y)$ = Local Lift Coefficient resulting from the induced angle of attack due to roll rate $\frac{py}{U}$.

Following [13], the equation of motion for torsion can be written as

$$\frac{d}{dy} \left(GJ \frac{\theta}{dy} \right) = -t(y) \quad (42)$$

where $t(y)$ is the applied external torque per unit span. Here we assume that the effect of smart material is essentially to produce an external torque and thus we ignore the contributions of material properties of the smart material related to torsion. The applied external torque per unit span is given by

$$t(y) = (ecC_L^e + c^2C_{m_{Ac}})q - Nmgd + t_{pz}(y) - qec\frac{\partial C_L^{ir}}{\partial \dot{\theta}}\left(\frac{p^l}{U}\right) \quad (43)$$

Here we approximate that the rate of change of lift coefficient due to change in the roll rate induced angle of attack is essentially constant and is given by its value at the wing tip. This way we need to determine this 'stability derivative' only once (albeit at the expense of some accuracy which presumably is tolerable). In the above expression, $t_{pz}(y)$ is the induced torque per unit span due to the smart material embedded in the wing. The boundary conditions for θ are $\theta(0) = \theta'(l) = 0$. The integral equation of torsional static equilibrium is obtained as

$$\theta(y) = \int_0^l C^{\theta\theta}(y, \eta)[(ecC_L^e + c^2C_{m_{Ac}})q - Nmgd + t_{pz}]d\eta \quad (44)$$

We now introduce 'generalized coordinates' and write all the spatial dependent variables in terms of these generalized coordinates. Note that since we are considering only the *static* problem, these generalized coordinates are *constants*. So let us write

$$\theta(y) = \sum_{j=1}^n \Theta_j(y)q_j \quad (45)$$

$$C_L^e(y) = \sum_{j=1}^n C_{L_j}^e(y)q_j \quad (46)$$

$$t_{pz}(y) = \sum_{j=1}^n \kappa_{pz}(y)q_j \quad (47)$$

Note that $\Theta_j(y)$ are the 'Structural Mode Shapes'. $C_{L_j}^e(y)$ are the 'Aerodynamic Mode Shapes'. $\kappa_{pz}(y)$ are the 'Piezoelectric Mode Shapes'. The Lagrange's Equation of Motion then gives

$$\frac{\partial U}{\partial q_i} = Q_i \quad (48)$$

where Q_i are the 'generalized forces'. The 'Strain Energy' can be written as

$$U = \frac{1}{2} \int_0^l GJ(\theta')^2 dy \quad (49)$$

which in turn can be written as

$$U = \frac{1}{2} \int_0^l GJ \left(\sum_{j=1}^n \Theta_j(y) q_j \right)^2 dy \quad (50)$$

We thus have the left hand side as

$$\frac{\partial U}{\partial q_i} = \sum_{j=1}^n \left(\int_0^l GJ \theta'_i \theta'_j dy \right) q_j \quad (i, j = 1, 2 \dots n) \quad (51)$$

The generalized force Q_i is obtained from

$$Q_i = \int_0^l t(y) \theta_i(y) dy \quad (52)$$

$$= \int_0^l (ec C_L^e + c^2 C_{m_{Ac}}) q - Nmgd + t_{pz}(y) + qec \frac{\partial C_L^{ir}}{\partial \frac{pl}{U}} \left(\frac{pl}{U} \right) \theta_i(y) dy \quad (53)$$

$$= \int_0^l [qec C_L^e + t_{pz} + t^*(y)] \theta_i(y) dy \quad (54)$$

$$(55)$$

where

$$t^*(y) = -Nmgd + t_{pz}(y) + qec \frac{\partial C_L^{ir}}{\partial \frac{pl}{U}} \left(\frac{pl}{U} \right) + qc^2 C_{m_{Ac}} \quad (56)$$

The equation of motion can be further written as

$$\sum_{j=1}^n \left[\int_0^l GJ \theta'_i \theta'_j dy - q \int_0^l C_{L_j}^e \theta_i cedy - \int_0^l \kappa_{pz}(y) \theta_i dy \right] = \int_0^l t^* \theta_i dy \quad (57)$$

We thus have

$$A_{ij} q_j = b_i \quad (58)$$

where

$$A_{ij} = \int_0^l [GJ \theta'_i \theta'_j - q C_{L_j}^e \theta_i ce - \kappa_{pz}(y) \theta_i] dy \quad (59)$$

$$b_i = \int_0^l t^* \theta_i dy \quad (60)$$

The boundary conditions of structural mode shape θ are:

$\theta(0) = 0, \theta(l)' = 0$. Thus we can finally solve for deformations q_j for given roll rate p (i.e. the right hand side vector) and for known Structural, Aerodynamic and Piezoelectric Mode shape information).

The above relationship connects the desired roll rate p with the required elastic twist distribution and the required smart material excitation.

Clearly, this is a preliminary step to understand the fundamental phenomenon and considerably more research is needed to obtain solutions of practical value.

6 Conclusions:

In this research, new efforts are undertaken to design controllers that explicitly take the nature of sensing (such as acceleration measurements) into account and also provide more flexibility in the use of various combination of sensors thereby providing improved alternatives to State Space design methods. Then integrated structure/control optimization is carried out to optimize weight and control performance in the presence of unmodeled dynamics and real parameter variations. Finally a preliminary research effort was undertaken to design control logic for the static aeroelastic roll maneuvering control of an aircraft using active deformations of smart wing structures.

6.1 Future Research That Needs AFOSR Support:

There is clearly a need to carry out further research in these important areas. Firstly more research should be undertaken to come up with improved control design algorithms, in the lines of present research proposed in this report. Secondly the integrated structure/ control optimization ideas need to be expanded to include other interactions such as fluid/structure and fluid/structure/control. Last but not least is the need to continue the analytical studies in the above areas, especially the flight maneuver control using smart material excitation. Of course, in addition to analytical studies, considerable effort should be expended in conducting application studies to realistic scenarios involving software and hardware issues for industrial structures such as real wings. Finally the importance of conducting experimental studies to validate the theories can not be overemphasized. This a long term commitment for research in " Flight Control with Smart Deformable Wing Structures" must be of high priority for Universities such as The Ohio State University and Air Force (WL) and Aircraft Industries to reap the benefits of this exciting research.

References

- [1] N.G. Creamer and J.L. Junkins. Low-authority eigenvalue placement for second-order structural systems. *jgcd*, 14:698–701, May 1991.
- [2] J-N Juang and P.G. Maghami. Robust eigensystem assignment for second order dynamic systems. *AIAA Structural Dynamics Specialist Conference*, 14:10–18, 1991.
- [3] A. Suleman and V.B. Venkayya. Flutter control of an adaptive composite panel. *Adaptive Structures Forum, AIAA-94-1744-CP*, pages 118–126, 1994.
- [4] S.J. Hollowell and J. Dugundji. Aeroelastic flutter and divergence of stiffness coupled, graphite/epoxy cantilevered plates. *Journal of Aircraft*, 21:69–76, jan 1984.
- [5] T.N. Dracopoulos and H. Oz. Integrated aeroelastic control optimization of laminated composite lifting surfaces. *Journal of Aircraft*, 29(2):280–288, mar 1992.
- [6] Anjali M. Diwekar and Rama K. Yedavalli. Smart structure control in second order form. *Journal of Smart Structures and Materials*, 5:429–436, 1996.
- [7] M.N. Nasr-Bismarck. Finite element analysis of aeroelasticity of plates and shells. *Applied Mechanics Reviews*, 45(12):461–482, dec 1992.
- [8] N.S.Khot and D.E. Veley. Robustness characteristics of optimum structural/control design. *Journal of Guidance, Control and Dynamics*, 15(1):81–87, Jan-Feb 1992.
- [9] N.S.Khot. Optimum structural design and robust active control using singular value constraints. *Computational Mechanics*, 16:208–215, 1995.
- [10] D.S.Bernstein and W.M. Haddad. Lqg control with an h-infinity performance bound: A riccati equation approach. *IEEE AC*, 34(3):293–305, March 1989.
- [11] K.M.Sobel, S.S.Banda, and H.H. Yeh. Robust control for linear systems with structured state space uncertainty. *IJC*, 50(5):1991–2004, 1989.
- [12] F.Eastep N.S.Khot. and R.Kolonay. A method for enhancement of the rolling maneuver of a flexible wing. *Proceedings of the AIAA SDM Conference*, Salt Lake City, 1996.

-
- A 3D diagram of a rectangular plate. The plate has a length L along the X -axis, a width $2b$ along the Y -axis, and a thickness c along the Z -axis. A coordinate system is shown with the Z -axis pointing upwards, the Y -axis pointing along the width, and the X -axis pointing along the length.

A 3D perspective diagram of a wing in a flow field. The wing is shown with a dashed line representing its reference axis (elastic axis). A coordinate system is established with the x-axis along the flow direction, the z-axis perpendicular to it, and the y-axis along the span. The flow velocity U is indicated by an arrow. The dynamic pressure is denoted by q . The line of aerodynamic centers is shown as a dashed line along the wing's length. The line of centers of gravity is shown as a solid line. The distance between these two lines is labeled e . The aerodynamic force is represented by a vector $c_m(AC) c^2 q$ acting perpendicular to the flow. The gravitational force is represented by a vector Nmg acting downwards. The distance from the reference axis to the center of gravity is labeled d . The wing's surface is shown with a hatched pattern.

69-20

A FOURTH-ORDER, TIME-DOMAIN ALGORITHM
FOR MAXWELL'S EQUATIONS

Jeffrey L. Young
Associate Professor
Department of Electrical Engineering

University of Idaho
Moscow, ID 83844-1023

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Laboratory, Flight Dynamics Directorate

August 1996

A FOURTH-ORDER, TIME-DOMAIN ALGORITHM FOR MAXWELL'S EQUATIONS

Jeffrey L. Young
Associate Professor
Department of Electrical Engineering
University of Idaho

Abstract

A compact, central-difference approximation, in conjunction with the Yee grid, is used to compute the spatial derivatives in Maxwell's equations. To advance the semi-discrete equations, the four-stage Runge-Kutta integrator is invoked. This combination of spatial and temporal differencing leads to a scheme that is fourth-order accurate, conditionally stable and highly efficient. Moreover, the use of compact differencing allows one to apply the compact operator in the vicinity of a perfect conductor – an attribute not found in other higher-order methods. Results are provided that quantify the spectral properties of the method. Simulations are conducted on problem spaces that span one and three dimensions and whose domains are of the open and closed type. Results from these simulations are compared with exact, closed-form solutions; the agreement between these results is consistent with numerical analysis.

A FOURTH-ORDER, TIME-DOMAIN ALGORITHM FOR MAXWELL'S EQUATIONS

Jeffrey L. Young
Associate Professor
Department of Electrical Engineering
University of Idaho

Summary of Achievements and Contributions

Below is a brief summary of the specific achievements and contributions made at the Wright Patterson Air Force Base, Flight Dynamics Directorate, Aeromechanics Division during the summer of 1996:

- A fourth-order, time-domain algorithm has been developed and analyzed for the simulation of propagation and scattering of electromagnetic waves. Spatial discretization was accomplished by using compact, central differences; a four-stage Runge-Kutta integrator was used for the time advancement.
- The algorithm has been implemented into FORTRAN code. Three simulations were conducted to validate the algorithm and code:
 1. Pulse propagation on a one-dimensional grid with periodic boundary conditions: The pulse maintained its spectral integrity after propagating 256 wavelengths.
 2. Pulse propagation on a one-dimensional cavity with perfectly conducting boundaries: After eighty reflections, the data was still stable and the pulse shape was still intact.
 3. Carrier-wave propagation in a three-dimensional rectangular waveguide: A continuous sinusoid was propagated over 200 wavelengths before dissipation and dispersion errors became apparent. Both TE and TM modes were examined.
- The paper entitled "Towards the construction of a fourth-order difference scheme for transient wave simulation: Staggered grid approach," by J. L. Young, D. Gaitonde and J. S. Shang was submitted to the *IEEE Transactions on Antennas and Propagation*.

Introduction

The classical central difference method of Yee [1] has been thoroughly tested and its capabilities have been established. Application problems ranging from radar cross-section analysis [2] to hyperthermia treatment planning [3] have been solved by using this simple, but robust, tool. For present day supercomputers, problem spaces that extend tens of wavelengths per dimension and run times that last on the order of several thousand time steps are now possible. Unfortunately, for electrically large domains and for late time analysis, the classical Yee scheme begins to show its limitations, due to the accumulation of phase errors.

In the genre of partial differential equation solvers, one possible way to overcome the phase error issue addressed previously is to consider a class of higher-order difference approximations. For example, in the work of Fang [4], a fourth-order, central difference approximation is postulated in conjunction with the

Yee grid and a modified fourth-order leap-frog integrator is employed. By analyzing its spectral properties, Shlager, *et al.* [5] have shown that Fang's scheme is more efficient than Yee's for a desired accuracy level. Unfortunately, the computation of the spatial derivatives near boundaries is difficult for fourth-order derivatives due to the length of the difference stencil.

To overcome the problems associated with the application of higher-order stencils near boundaries, Hadi, *et al.* [6] have proposed a modified two-four FDTD scheme. In this work, the standard second-order, central approximations are used near perfect conductors while higher-order accuracy is achieved by incorporating tuning parameters in the temporal integration equations at nodes away from the boundary. For periodic domains, their dispersion analysis indicates that structures on the size of 1000 wavelengths on a side can be modeled with as little as six points per wavelength. However, it is not clear at this time what the impact of lowering the accuracy by two at the vicinity of a boundary will be on the overall accuracy of the scheme. Gustafsson [7, 8] suggests that such practices will create schemes that are closer to third-order.

With respect to nodally collocated schemes, higher-order approximations have been postulated and tested. One popular methodology is the MUSCL scheme in conjunction with Runge-Kutta integration [9]. Here, the user has the flexibility to select from a wide variety of approximations, with the third-order approximation being the most popular. Except for its dissipational properties and the ambiguity of specifying boundary values at perfect conductors to all six components of Maxwell's equations, the MUSCL scheme has shown great promise in solving a wide range of problems. In addition to accuracy issues, the MUSCL scheme is straightforward to implement on structured, non-Cartesian grids.

In this paper, we consider an efficient fourth-order in time and space approximation for the derivatives that appear in Maxwell's equations. By retaining the Yee stencil, we construct a compact difference operator for the spatial derivatives and use Runge-Kutta integration for the temporal advancement. (Note: In the work of Cangellaris [10] and others, the term compact is invoked to denote the replacement of a spatial derivative with its spectral representation. In our usage, compact signifies the relative size of the stencil.) The three key advantages of this approach are: 1) the boundary conditions are satisfied point-wise exactly, 2) the computational efficiency is dramatically improved (with respect to Yee's and Fang's scheme), and 3) the numerical algorithm retains much of the original simplicity of the Yee scheme. The main disadvantage is that the four-stage Runge-Kutta integrator requires additional memory (with respect to the Yee scheme) to advance the discrete equations one time step.

To demonstrate the validity of the proposed method, several examples are presented – these problems range from domains that are one-dimensional to three-dimensional. The first case is the problem of gaussian pulse propagation on a one-dimensional grid with periodic boundary conditions. This problem serves to exemplify the superb phase and dissipation error characteristics of the method. The example of a one-dimensional cavity is considered next as a forum for discussion on the impact of perfectly conducting boundaries on the solution. Finally, the propagation of a single frequency wave in a rectangular waveguide is simulated. This three-dimensional problem is considered due to the availability of a simple closed-form solution. In addition, the rectangular waveguide serves to show that the boundary conditions and the mesh truncation conditions can be implemented in three-dimensional spaces without encountering numerical instabilities or poor phase error characteristics.

In a companion paper [11], Gaitonde *et al.* discusses in detail how a nodally collocated finite-volume compact operator can be constructed to solve these same problems. In that paper the ease with which the finite-volume scheme can be cast in generalized coordinates is demonstrated by considering scattering from

perfectly conducting objects.

Formulation

For simple, lossless media and source-free conditions, the space-time relationship between the electric field vector \mathbf{E} and the magnetic field intensity \mathbf{H} is completely described by Maxwell's equations:

$$\epsilon \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H} \quad (1)$$

and

$$\mu \frac{\partial \mathbf{H}}{\partial t} = -\nabla \times \mathbf{E}, \quad (2)$$

with the stipulation that $\nabla \cdot \mathbf{E}$ and $\nabla \cdot \mathbf{H}$ are identically zero. The *a priori* assumption that the medium is simple requires that the permittivity ϵ and the permeability μ be piecewise constant, linear, isotropic, time-invariant and non-dispersive. In this paper, ϵ and μ will take on free space values and, for the numerical validation, they will be normalized to unity.

Spatial Differences

Given the natural interaction between field constituents and the implicit satisfaction of the divergence relationships, the Yee grid is invoked for the subsequently described numerical procedure. As a result, a class of dissipationless central-difference operators may be designed for the spatial derivatives that are either second- or fourth-order accurate. For example, the classical second-order central difference operator is simply

$$\left. \frac{\partial f}{\partial x} \right|_i \approx \frac{f_{i+1/2} - f_{i-1/2}}{\Delta_x} - \frac{\Delta_x^2 f^{(3)}(x)}{24}. \quad (3)$$

Here f is regarded as any one component of either \mathbf{E} or \mathbf{H} and Δ_x is the cell size in the x -direction. One principal feature of this operator is discovered when the boundary conditions for \mathbf{E} and \mathbf{H} at a perfect electrical conductor (PEC) are enforced. Since a perfectly conducting boundary is defined where tangential \mathbf{E} and normal \mathbf{H} vanish, one can enforce these conditions exactly and compute all other components near the boundary, without any additional approximations, information or equations.

For fourth-order accuracy, one may use the approximation of Fang:

$$\left. \frac{\partial f}{\partial x} \right|_i \approx \frac{-f_{i+3/2} + 27f_{i+1/2} - 27f_{i-1/2} + f_{i-3/2}}{24\Delta_x} + \frac{3\Delta_x^4 f^{(5)}(x)}{640} + \frac{\Delta_x^6 f^{(7)}(x)}{3584}. \quad (4)$$

It is readily seen that the stencil of this operator extends to two nodes on either side of the unknown at x . Consequently, there appears to be no clear way to apply this operator directly in the vicinity of a perfectly conducting boundary or with domain truncation operators (such as Mur's [12]).

An efficient way to limit the size of the stencil is to consider the class of compact operators [13]. For the case at hand, we consider the following compact central operator for the staggered grid:

$$\alpha_1 \left. \frac{\partial f}{\partial x} \right|_{i+1} + \alpha_2 \left. \frac{\partial f}{\partial x} \right|_i + \alpha_1 \left. \frac{\partial f}{\partial x} \right|_{i-1} = \beta \frac{f_{i+1/2} - f_{i-1/2}}{\Delta_x}, \quad (5)$$

where the unknown coefficients $\alpha_1, \alpha_2, \beta$ are to be determined. Since the compact operator is implicit in nature, the determination of the derivative of f at x is globally connected to the derivative of f at other

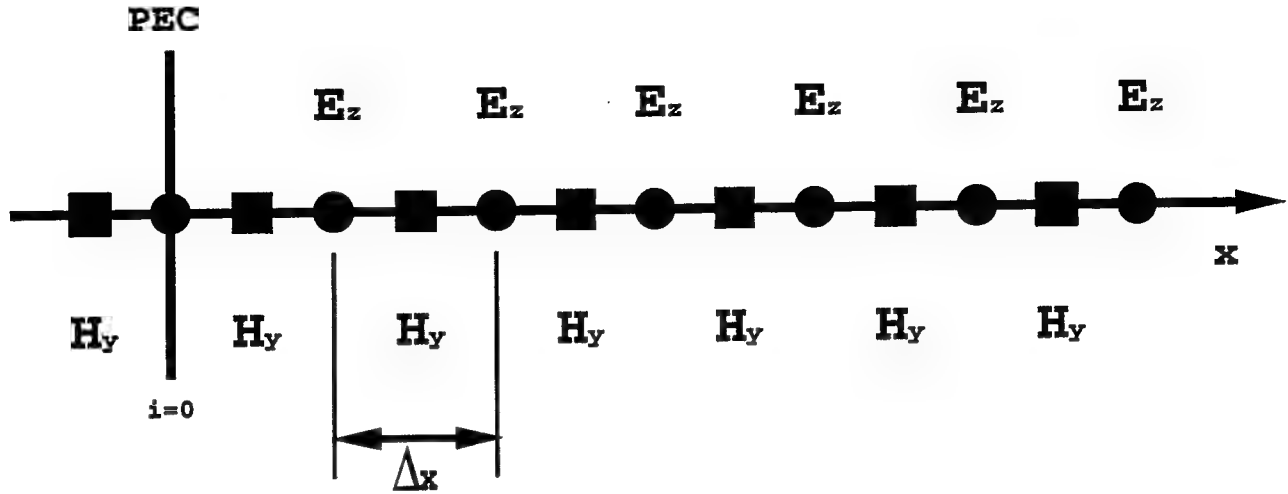


Figure 1: A one-dimensional depiction for the staggered grid truncated by a perfect electrical conductor (PEC).

points in the domain; the connectivity of derivatives is established when the matrix of the previous system of equations is inverted. However, the structure of the linear system of equations is tridiagonal with the implication that the solution for the derivative of f at x is easily determined via the efficient Thomas algorithm [14], which requires only 5 additional operations per derivative (i.e., $5N$ operations are required to invert the tridiagonal matrix, where N is the matrix size.) Since the right-hand side of Eqn. (5) is similar in structure to the right-hand side of Eqn. (3), the boundary conditions at perfectly conducting boundaries are treated no differently than Yee's scheme. That is, this compact operator retains an important feature of the second-order central operator by satisfying the boundary conditions point-wise exactly.

To apply this compact operator at or near a perfectly conducting interface, consider Figure 1, which shows a one-dimensional depiction of the Yee cell. Under the assumption that the PEC is at $x = 0$, we observe that the computation of the derivative of H_y with respect to x at $x = \Delta_x$ requires knowledge of the derivative of H_y at $x = 0$, $x = \Delta_x$ and $x = 2\Delta_x$. Due to the construction of the Yee cell, $E_y = E_z = H_x = 0$ at $x = 0$, which requires that $\partial H_y / \partial x$ be zero also. Hence, the evaluation of the compact difference for the derivative of H_y with respect to x requires no special treatment near or at a boundary. This statement also holds true for all other spatial derivatives of \mathbf{H} that are associated with Ampere's law.

The spatial derivatives of \mathbf{E} associated with Faraday's law are not so trivial to compute. Again, from Figure 1 we observe that to compute the derivative of E_z with respect to x at $x = \Delta_x/2$ we must have knowledge of this derivative at $x = -\Delta_x/2$, $x = \Delta_x/2$ and $x = 3\Delta_x/2$. Unfortunately, the point $x = -\Delta_x/2$ lies outside the computational domain. To calculate an edge value for the Thomas algorithm for the derivative at $x = \Delta_x/2$, an explicit, central biased, 6th-order difference is used:

$$\left. \frac{\partial f}{\partial x} \right|_{i+1/2} = [b_1 f_i + b_2 f_{i+1} + b_3 f_{i+2} + b_4 f_{i+3} + b_5 f_{i+4} + b_6 f_{i+5} + b_7 f_{i+6}] / (1920 \Delta_x), \quad (6)$$

where $b_1 = -1627$, $b_2 = 633$, $b_3 = 2360$, $b_4 = -2350$, $b_5 = 1365$, $b_6 = -443$ and $b_7 = 62$. The equivalence

between f and E_z requires that f be forced to zero at $x = 0$. The reasons for choosing a sixth-order operator are subjective: 1) Using a sixth-order approximation with a reasonably small Δ_x , we force the truncation term to be lower than the truncation term for the interior region. Hence, the accuracy of the overall scheme will be highly dependent upon the compact difference operator rather than upon the edge value. 2) It is quite possible that numerical instabilities could arise due to the edge value for the Thomas algorithm (numerical experimentation so far reveals that this is not the case). The sixth-order approximation will force the instabilities (if any) to be associated with the very high frequencies and these frequencies may be damped sufficiently by the fourth-order temporal integrator, which is known to be relatively dissipative at large time steps.

To determine the values of the coefficients, consider the following Fourier analysis. Assume that f is of the form $e^{jkx-j\omega t}$, which describes a plane wave propagating in the positive x direction. When this wave is discretized using Eqn. (5), we obtain

$$K(q) = \frac{\beta(2/\Delta_x) \sin(q/2)}{2\alpha_1 \cos(q) + \alpha_2}, \quad (7)$$

where K is the numerical wavenumber and $q = k\Delta_x$. Ideally, the numerical operator should produce a K/k (i.e., normalized wavenumber) value of unity for all wavenumbers. Practically, K/k is unity at only one wavenumber, and close to unity over a limited range of spatial frequencies. Typically, the wavenumber of choice for which K/k takes on a value of unity is zero. For example, if Taylor's theorem is applied to K about the point $q = 0$, like terms are equated and the second-order term is forced to zero, the standard compact, fourth-order, central difference operator is constructed; for this case, $\alpha_1 = 1$, $\alpha_2 = 22$ and $\beta = 24$. Of course, when $\alpha_1 = 0$, $\alpha_2 = 1$ and $\beta = 1$, the compact operator reverts to a classical, explicit, second-order, central operator. (Other choices for α_1 , α_2 and β may be selected based upon various optimization criteria; this will be a subject of future study.)

With $\alpha_1 = 1$, $\alpha_2 = 22$ and $\beta = 24$, an equivalent expression for (5) that reveals the truncation terms is

$$\left. \frac{\partial f}{\partial x} \right|_i \approx \frac{24\delta f_i}{\Delta_x(24I + \delta^2)} + \frac{17\Delta_x^4 f^{(5)}(x)}{5280} + \frac{109\Delta_x^6 f^{(7)}(x)}{887040}, \quad (8)$$

where $\delta f_i = f_{i+1/2} - f_{i-1/2}$ and I is the identity operator. By comparing Eqn. (4) with Eqn. (8), we observe that both the fourth- and sixth-order truncation terms for the compact operator are lower than those for the fourth-order operator of Fang. Particularly, the ratio (compact to explicit) of the leading fourth-order coefficients is 0.69; the ratio of the sixth-order coefficient terms is 0.44. This rapidly convergent series for the compact operator gives rise to a highly accurate difference scheme.

Some of the previous concepts are demonstrated by considering Figure 2. In this figure we plot K/k as a function of $k\Delta_x$ for the second-order, fourth-order, and compact fourth-order operators. Clearly, the second-order central operator has limited spectral bandwidth when compared to its fourth-order counterparts. In addition, the slight increase in the spectral bandwidth of the compact operator as compared to Fang's operator is attributable to its low leading truncation term. To quantify these previous statements, we define the bandwidth of each of these operators as the highest spatial frequency for which K/k does not drop below a value of 0.999. From the data, the second-order operator has a bandwidth of $k\Delta_x = 0.154$; for the fourth-order, $k\Delta_x = 0.684$; for the fourth-order compact, $k\Delta_x = 0.762$. Thus, for a phase error of 0.001 r/m, the wave should be sampled at 40.8 cells per wavelength when using the second-order operator, 9.2 cells per wavelength when using the fourth-order operator, and 8.2 cells per wavelength when using the fourth-order compact operator.

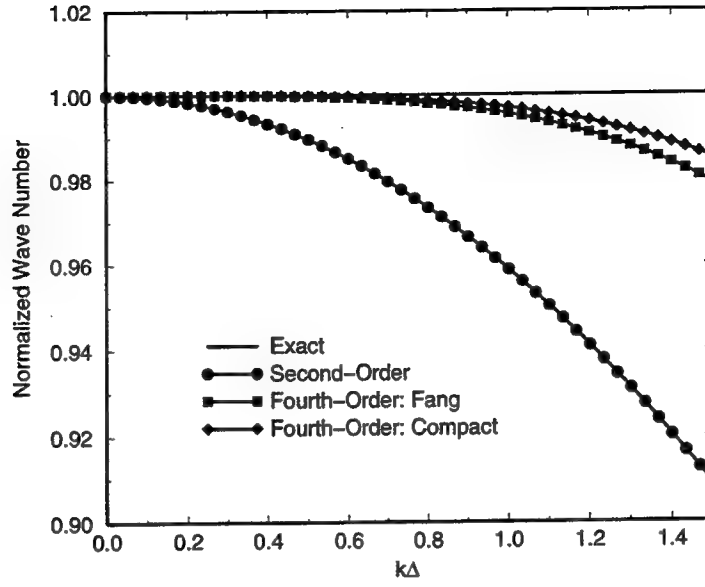


Figure 2: Phase characteristics of three numerical operators as a function of normalized spatial frequency.

Temporal Integration

To achieve fourth-order temporal accuracy, there exist several standard possibilities, two of which are considered. First consider the fourth-order, leap-frog method of Fang. In this approach, the value of f at $t = (n + 1)\Delta_t$ is determined from f at $n\Delta_t$, a first-order derivative at $(n + 1/2)\Delta_t$ and a third-order correctional derivative at $(n + 1/2)\Delta_t$:

$$f^{n+1} = f^n + \Delta_t \left. \frac{\partial f(x, t)}{\partial t} \right|^{n+1/2} + \frac{\Delta_t^2}{24} \left. \frac{\partial^3 f}{\partial t^3} \right|^{n+1/2}. \quad (9)$$

However, through repeated application of Maxwell's equations this correctional derivative is converted into spatial derivatives. For example, when integrating Ampere's law, Fang reported that

$$\epsilon \mathbf{E}^{n+1} = \epsilon \mathbf{E}^n + \Delta_t \nabla \times \mathbf{H}^{n+1/2} + c^2 \frac{\Delta_t^3}{24} \nabla \times (\nabla^2 \mathbf{H}^{n+1/2}). \quad (10)$$

The curl-Laplacian term is computed to second-order since Δ_t is related to Δ_x via the CFL number. Fang's scheme is dissipationless and requires no back storage of the dependent variables. However, the incorporation of boundary and truncation conditions is still problematic and the computation of the correctional curl-Laplacian term is neither efficient nor trivial.

We employ the class of Runge-Kutta (RK) integrators in this work. The two-stage RK scheme is unconditionally unstable when the spatial difference operators admit eigen-values that are purely imaginary. Such is the case with both the explicit and compact central difference operators. However, the four-stage RK scheme will remain conditionally stable, as specified by the CFL number, even for central differences. Formally, the RK methods may be expressed in terms of a truncated Taylor series for the matrix $e^{-A\Delta_t}$,

where A is the spatial discretization matrix:

$$f^{n+1} = \sum_{m=1}^M \frac{(-\Delta_t A)^m}{m!} \cdot f^n. \quad (11)$$

Obviously, the implementation of this series is straightforward and any order of accuracy can be achieved, if one is willing to expend the computational resources to achieve that accuracy. However, to minimize the number of evaluations of the right-hand side of Eqn. (11) and to limit the storage required for each evaluation, a fourth-order scheme is used [15]. That is,

$$\begin{aligned} k_0 &= \Delta_t R(f_0) \\ k_1 &= \Delta_t R(f_1) \\ k_2 &= \Delta_t R(f_2) \\ k_3 &= \Delta_t R(f_3) \\ f^{n+1} &= f^n + (k_1 + 2k_2 + 2k_3 + k_4)/6. \end{aligned} \quad (12)$$

Here R is the residual; $f_0 = f(x, t_0)$, $f_1 = f_0 + k_0/2$, $f_2 = f_1 + k_1/2$ and $f_3 = f_2 + k_2$.

One property of the RK integrators is the admittance of numerical dissipation and dispersion into the solution. This is best seen by considering the fully discretized system of equations in terms of a complex amplification factor σ , which is defined by the ratio f^{n+1}/f^n . In spectral space, the fields are assumed to be time-harmonic plane waves propagating through a uniform, infinite, one-dimensional grid. Based on this assumption and the model wave equation [16],

$$\sigma = 1 - j\nu K \Delta_x - (\nu K \Delta_x)^2/2 + j(\nu K \Delta_x)^3/6 + (\nu K \Delta_x)^4/24, \quad (13)$$

where ν is the CFL number; $\nu = c\Delta_t/\Delta_x$. For comparison purposes, the amplification factor for the Yee scheme is

$$\sigma = 1e^{-j\phi}, \quad (14)$$

where

$$\phi = -2 \sin^{-1} \left[\nu \sin \left(\frac{k\Delta_x}{2} \right) \right]. \quad (15)$$

For Fang's scheme, the amplitude factor is also unity and the phase term is given by [5, 17],

$$\phi = -2 \sin^{-1} \left\{ \nu \sin \left(\frac{k\Delta_x}{2} \right) \left[1 + \frac{(1-\nu^2)}{6} \sin^2 \left(\frac{k\Delta_x}{2} \right) \right] \right\}. \quad (16)$$

Numerical instability is encountered when the magnitude of σ exceeds the value of unity for any frequency. By employing a numerical search on Eqn. (13), we found that a ν of 1.1784 is the maximum CFL value for the RK-compact scheme. Numerical experiments confirm this value. Yee's and Fang's scheme remain stable for CFL values up to unity.

For a CFL value of unity, the previous three equations indicate that the Yee and Fang schemes admit the perfect shift condition in one-dimensional spaces. In three-dimensional spaces, this ideal is not achieved for all wavenumbers. To make a comparison study between algorithms, consider a three-dimensional space that is discretized with cubic cells. According to stability theory, the time step is set by the CFL number using the rule,

$$\Delta_t = \frac{\nu \Delta_x}{\sqrt{3}c} = \frac{\nu' \Delta_x}{c}, \quad (17)$$

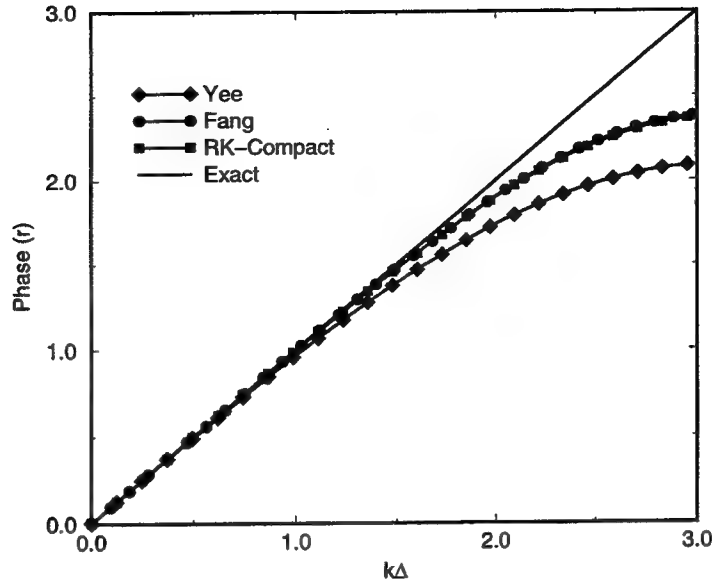


Figure 3: Dispersion data for three schemes: RK-compact, Fang and Yee; $\nu' = 0.5$.

where $\nu' = \nu/\sqrt{3}$. By assuming that the wave is traveling in the x -direction and by setting ν to equal ν' in the previous amplification equations, we can use those amplification equations to characterize the phase and amplitude performance of each scheme. When ν' is set to 0.5 for each scheme, the phase data shown in Figure 3 is obtained. As expected, both fourth-order schemes have nearly identical phase characteristics while the Yee scheme is highly dispersive. As an example, when $k\Delta_x = \pi/4$ (i.e., eight points per wavelength), the phase error for Yee's scheme is 0.0153 r; for Fang's scheme, the error is 0.00101 r; for the RK-compact scheme the error is 0.00103 r. Conversely, for Yee's scheme to achieve a phase error of 0.001 r, we must set $k\Delta_x$ to a value of 0.316 (i.e., 19.9 points per wavelength).

Dissipation data is shown in Figure 4. Although the compact spatial operator is dissipationless, the dissipation seen in Figure 4 is attributable to the RK integrator. For example, when the wave is sampled at eight points per wavelength the amplitude error is a value of 0.025%.

Algorithm Performance

On a uniform grid (i.e., $\Delta_x = \Delta_y = \Delta_z$) with no boundaries, 264 operations per cell per time step are needed to compute the three-dimensional electromagnetic field using fourth-order, RK-compact scheme (see [15] for one possible implementation of the fourth-order RK procedure.). In contrast, Yee's scheme requires 30 and Fang's requires 312 [5]. Based upon Shlager's et al. analyses, Fang's scheme is about 6.4 times more efficient than Yee's for a given phase error. Hence, it follows that the RK-compact is about 7.6 times more efficient than Yee's scheme and about 1.2 times as efficient as Fang's.

This twenty percent increase in performance of the RK-compact scheme over Fang's scheme is achieved at the expense of computational memory. That is, Fang's scheme manipulates six unknowns per cell; the

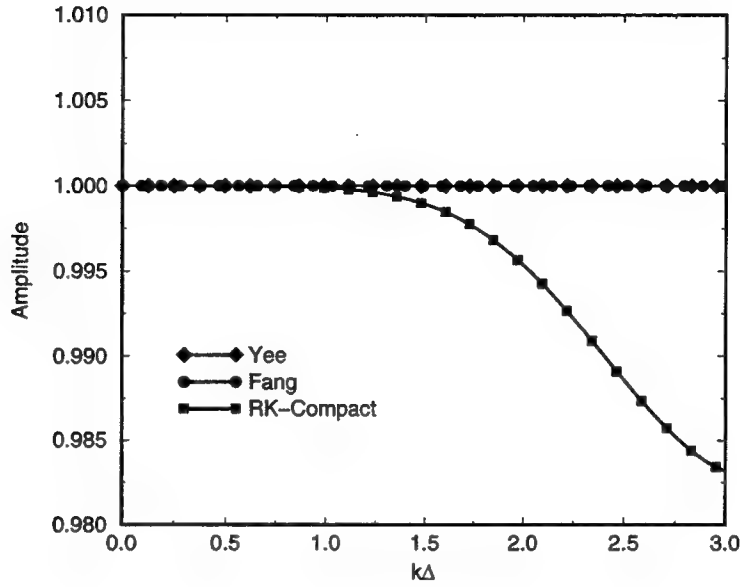


Figure 4: Dissipation data for three schemes: RK-compact, Fang and Yee; $\nu' = 0.5$.

RK-compact [15] manipulates thirty. Some memory reduction can be achieved by recognizing that the semi-discrete, time-dependent equations constitute a set of linear, constant-coefficient, ordinary differential equations. This property allows the fourth-order RK method to be reformulated so that only 18 [18] or 12 [19] unknowns per cell are manipulated. Note: Low-storage RK schemes may require more operations per time step, thus compromising the scheme's efficiency.

Numerical Results

The problem of a pulse propagating in a one-dimensional grid with periodic boundary conditions is studied first. The pulse is mathematically given by

$$f(t, z) = e^{-u^2 w^2}, \quad (18)$$

where $u = t - z$ and w , the pulse width parameter, is set to a value of 1.035. This w -value will insure that the pulse will contain sufficient spectral information to a frequency of π radians. At π radians, the amplitude of that spectral line is down to 0.1 of the dc value; at 2π radians, it is down to 0.0001 of the dc value. To capture the 2π angular frequency, the wave is sampled at 10 points/ λ when $\omega = 2\pi$ r. For this simulation, a CFL value of unity is chosen.

The time-domain error (i.e., the difference between exact and computed data) is displayed in Figure 5 after the pulse has traveled 5120 cells; or a length of 512λ at the 2π r spectral component. From the time-domain data, we observe that the maximum error is about 2.1%. To appreciate the effects of dissipation errors, Figure 6 is provided. Here we plot the amplitude of the Fourier response for both the exact and the numerically computed solutions. Note that the effect of numerical dissipation (a property of the RK

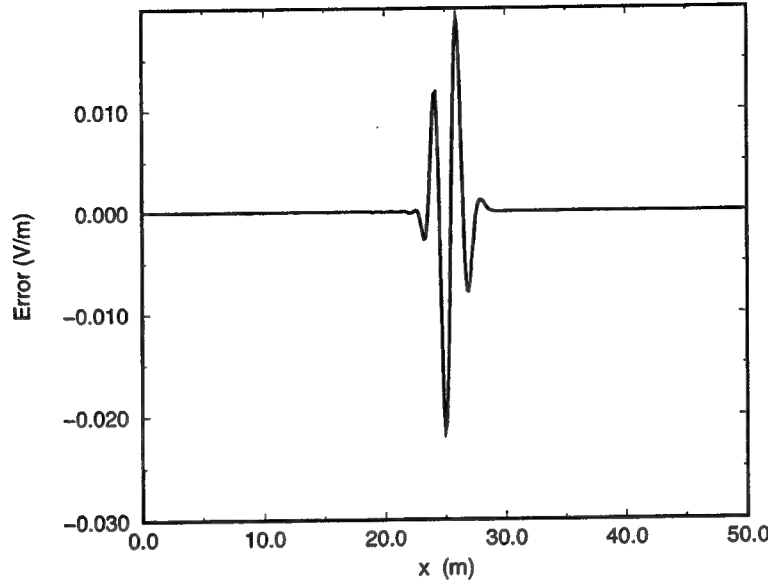


Figure 5: Pulse error after 5120 time steps: $\omega = 1.03$ 1/s, CFL=1.0, periodic boundary conditions.

integrator) is becoming evident at $\omega = \pi$ r/s.

To appreciate the effects of perfect conductors on the algorithm's performance, the one-dimensional cavity is studied next. Instead of incorporating periodic conditions at the domain's boundary, we truncate the domain with perfect electric conductors; all other features of the previous problem are still retained. For this situation, there is no evidence of any instabilities after 80 reflections, as seen in Figure 7. The plot in this figure shows the pulse after it has traveled 40,960 cells or a length of 4096λ at the 2π r spectral component. The effects of numerical dissipation and dispersion on the pulse's wave shape are obvious.

Finally, consider the problem of the rectangular waveguide. This three-dimensional problem is considered for three reasons: 1) a closed-form solution is available, 2) a mesh truncation condition must be imposed, and 3) boundary conditions at perfect conductors must be satisfied. In addition, the waveguide can be viewed as supporting a collection of plane waves constructively and destructively interfering with each other as they reflect from the perfectly conducting walls. To appreciate this latter point, consider the TE field component H_z . In the frequency domain,

$$H_z = H_0 \cos(k_x x) \cos(k_y y) e^{-j k_z z} \quad (19)$$

where $k_x = m\pi/a$, $k_y = n\pi/b$ and $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$. Or, equivalently,

$$H_z = (H_0/4) \left(e^{j(k_x x + k_y y - k_z z)} + e^{j(k_x x - k_y y - k_z z)} + e^{-j(k_x x - k_y y + k_z z)} + e^{-j(k_x x + k_y y + k_z z)} \right) \quad (20)$$

By letting $k_x = k \cos \theta_x$, $k_y = k \cos \theta_y$ and $k_z = k \cos \theta_z$, we complete the interpretation of four plane waves reflecting from the waveguide walls.

To initialize the fields, the exact, time-harmonic solution [20] is imposed everywhere on the six electromagnetic field components. Once the simulation has begun, the fields are driven at the plane defined

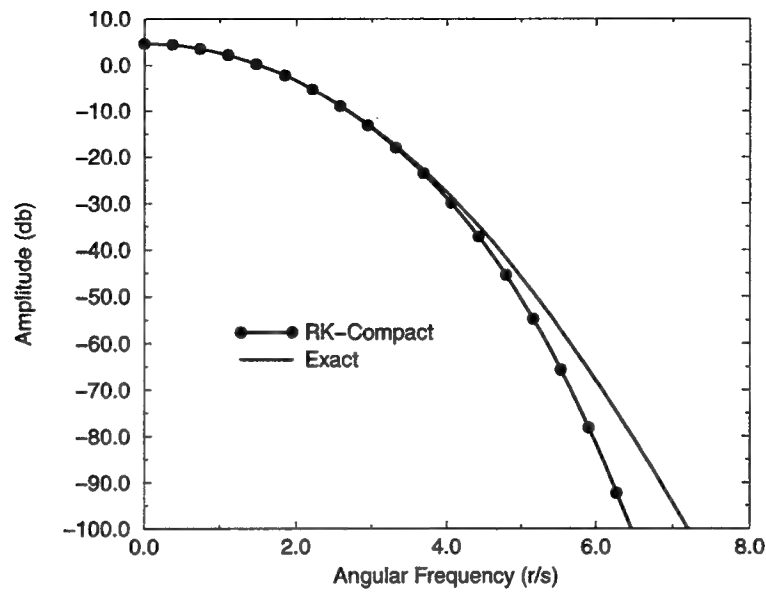


Figure 6: Amplitude response of the exact and numerically computed solutions after 5120 time steps: $w = 1.03$ 1/s, CFL=1.0, periodic boundary conditions.

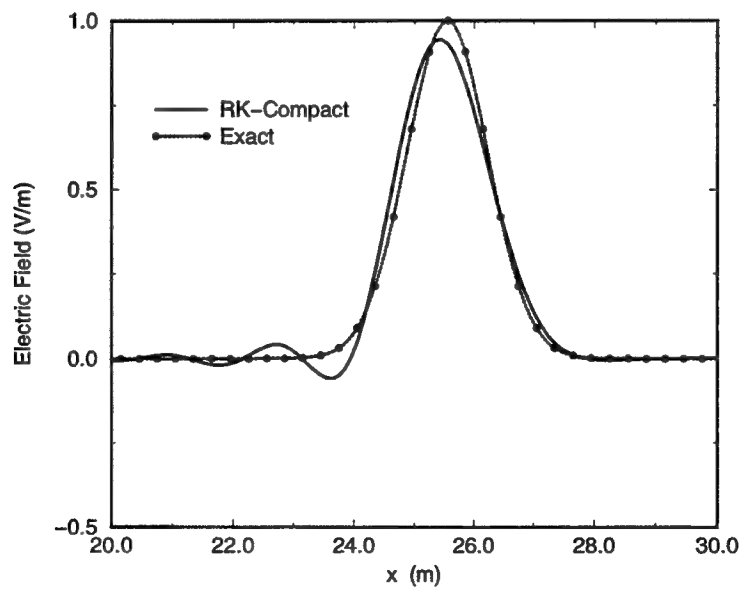


Figure 7: Time-domain response of the exact and computed solutions for the one-dimensional cavity: $w = 1.03$ 1/s, CFL=1.0, 40,960 time steps.

at $z = 0$. To truncate the domain at $z = l$, we found that the standard characteristic-based condition did not provide the accuracy required for this simulation. The deficiency is due to its inability to suppress adequately spurious reflections for angles near grazing. To circumvent this problem, a modal operator is used: For example, for H_y , we require

$$\frac{\partial H_y}{\partial t} + v_p \frac{\partial H_y}{\partial z} = 0, \quad (21)$$

where v_p is the phase velocity of the mode under consideration. This operator may be regarded as exact for any one mode and frequency. For multi-mode truncation, it is subject to many of the same problems as the standard characteristic-based condition. On the tz -plane, the boundary operator of Eqn. (21) can be viewed in terms of the model one-dimensional wave equation. From this perspective, characteristic theory suggests that the most stable way to implement the z -differentiation is to use windward differencing. For fourth-order accuracy, this translates into the following approximation:

$$\left. \frac{\partial f}{\partial z} \right|_i = [25f_i - 48f_{i-1} + 36f_{i-2} - 16f_{i-3} + 3f_{i-4}] / (12\Delta_z). \quad (22)$$

For the present study, a waveguide that is 10 cells by 10 cells in the cross-sectional plane and 2000 cells in length is chosen; cubic cells are chosen that are .1 m on a side. With $\omega = 2\pi$ r/s, the wave is sampled at 10 cells per free-space wavelength. For the TE_{11} and the TM_{11} modes, the effective longitudinal sampling is 14.14 cells per wavelength. With $2\Delta_t = \Delta_x$, 5600 time steps are needed for the wave to completely travel down the guide.

In Figures 8 and 9 we show the field H_z as a function of z for the TE_{11} along the axis of the waveguide. Figure 8 considers the field in the vicinity of $z = 1000\Delta_z$ (i.e., 100 free space wavelengths); Figure 9 considers the field in the vicinity of $z = 2000\Delta_z$ (i.e., 200 free space wavelengths). Even though both of these plots show the anticipated gradual build-up of dissipation and dispersion type errors, the scheme is able to provide stable and reasonably accurate data. Moreover, the data reveals that the waves do not encounter any noticeable artificial boundaries that induce reflections at the end of the guide.

This same experiment is repeated for the TM_{11} mode, except now we focus our attention on E_z . Consider Figures 10 and 11, which show the field in the vicinity of $z = 1000\Delta_z$ and $z = 2000\Delta_z$, respectively. Again, the phase and attenuation errors become more pronounced as the wave travels farther down the guide.

Conclusions

An efficient, fourth-order, time-dependent difference scheme has been developed and applied to electromagnetic wave problems. The problems chosen were of the open and closed type, and their geometries spanned one and three dimensions. By using a compact difference method, we were able to construct a fourth-order operator that can be used in the vicinity of a perfect conductor without any apparent limitation on the stability or accuracy of the scheme. A four-stage Runge-Kutta integration technique was employed to advance the semi-discrete equations. Although the traditional 4th-order RK integrator is highly efficient, the efficiency is achieved at the expense of computational memory. Conversely, further reduction in memory is possible if one is willing to sacrifice computational efficiency.

As our numerical analysis anticipated, the numerical results were eventually contaminated by numerical phase and attenuation errors. Nevertheless, these errors did not become significant until the wave had traveled well beyond 200 wavelengths in a three-dimensional waveguide. In one-dimensional spaces, we were able to propagate a pulse over 500 wavelengths without impacting the structural integrity of the pulse.

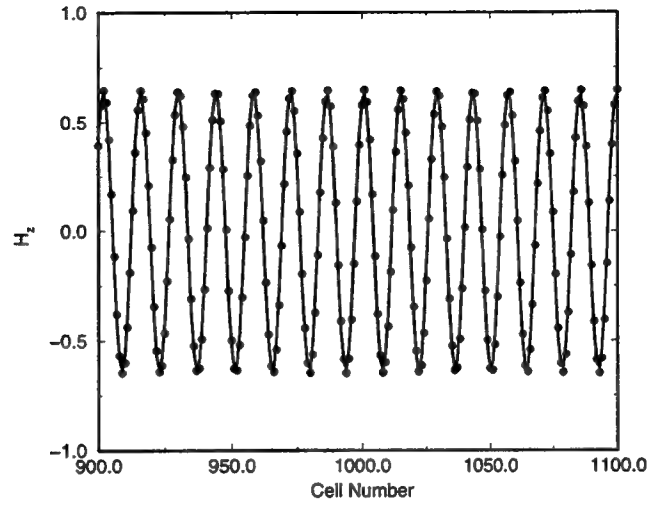


Figure 8: H_z as a function of z in a rectangular waveguide for the TE_{11} mode. The waveguide is 2000 cells in length; the simulation time is 5600 time steps; $2\Delta_t = \Delta_x = .1$; the wave is sampled at 10 points per free-space wavelength.

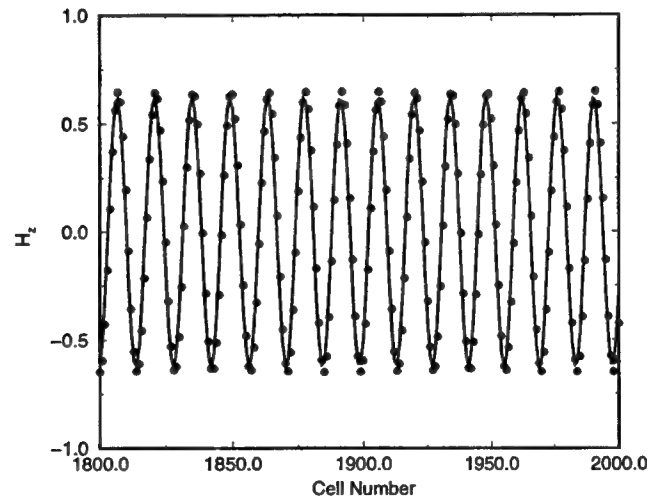


Figure 9: H_z as a function of z in a rectangular waveguide for the TE_{11} mode. The waveguide is 2000 cells in length; the simulation time is 5600 time steps; $2\Delta_t = \Delta_x = .1$; the wave is sampled at 10 points per free-space wavelength.

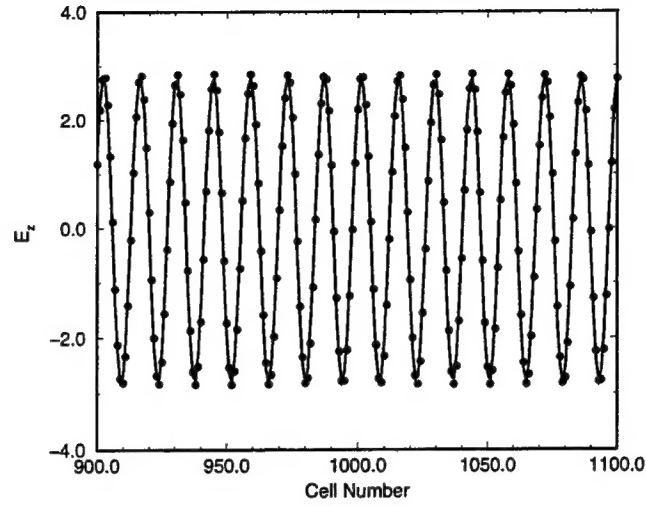


Figure 10: E_z as a function of z in a rectangular waveguide for the TM_{11} mode. The waveguide is 2000 cells in length; the simulation time is 5600 time steps; $2\Delta_t = \Delta_x = .1$; the wave is sampled at 10 points per free-space wavelength.

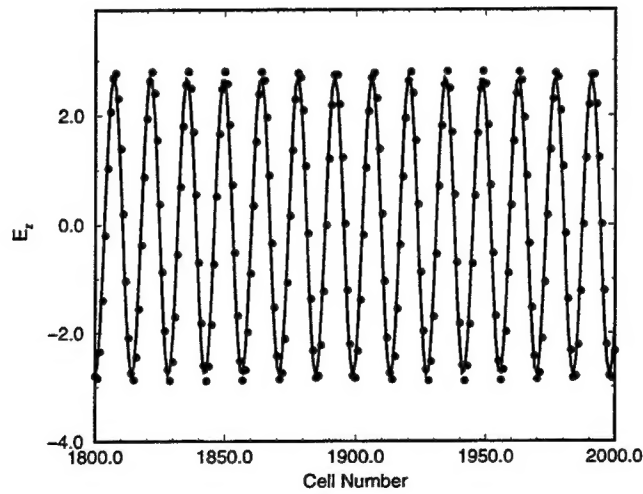


Figure 11: E_z as a function of z in a rectangular waveguide for the TM_{11} mode. The waveguide is 2000 cells in length; the simulation time is 5600 time steps; $2\Delta_t = \Delta_x = .1$; the wave is sampled at 10 points per free-space wavelength.

The RK-compact scheme retains much of the original simplicity of Yee's scheme. However, there is an added level of complexity in the RK-compact scheme that is not present in Yee's. Although boundary information for the dependent variables is incorporated in the exact same way as in Yee's method (i.e., the right hand side of Eqn. (5)), this same information must also be made available for the computation of the derivatives of the dependent variables (i.e., the left hand side of Eqn. (5)). For objects that are complex in shape, incorporating such information into code may represent a formidable task. Conversely, finite-volume solvers that use compact methods [11] in the reconstruction of fluxes at flux walls are naturally formulated to handle irregular geometries.

The domain truncation condition employed herein was devised for a specific problem. It is surmised that for scattering applications, a general fourth-order type ABC will be needed if unwanted spurious reflections are to be suppressed. Further study on the application, stability and accuracy of higher-order ABC's is being conducted.

Finally, due to the simplicity of the RK integrator and the Thomas algorithm, vectorization of the code is straightforward. With compiler optimization, the results for the waveguide problem were obtained in about 2.5 cpu hours on a Cray C90 machine. (Recall that the problem consisted of manipulating 1.2 million unknowns 5600 times). Given the structure of the code, a parallelized version is equally possible.

Acknowledgements

Supercomputing resources were provided by the DoD HPC Shared Resource Center, CEWES, Vicksburg, MS.

References

- [1] Yee, K. S., "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Ant. Propagat.*, vol. 14, pp. 302-307, 1966.
- [2] Taflove, A. and K. Umashankar, "Radar cross section of general three-dimensional scatterers", *IEEE Trans. Electromagn. Compat.*, vol. 25, pp. 433-440, 1983.
- [3] Sullivan, D. M., D. T. Borup and O. P. Gandhi, "Use of the finite-difference time-domain method in calculating EM absorption in human tissues," *IEEE Trans. Biomed. Eng.*, vol. 34, pp. 148-157, 1987.
- [4] Fang, J., *Time Domain Finite Difference Computation for Maxwell's Equations*, Ph.D Dissertation, University of California, Berkeley, 1989.
- [5] Shlager, K. L., J. G. Maloney, S. L. Ray, A. F. Peterson, "Relative accuracy of several finite-difference time-domain methods in two and three-dimensions," *IEEE Trans. Ant. Propagat.*, vol. 41, pp. 1732-1737, 1993.
- [6] Hadi, M. F., and M. Picket-May, "A modified FDTD (2,4) scheme for modeling electrically large structures with high phase accuracy," *12 Annual Rev. of Progress in Appl. Comp. Electromagn.*, Monterrey, CA, pp. 1023-1030, 1996.
- [7] Gustafsson, B., "The convergence rate for difference approximations to mixed initial boundary value problems," *Math. Comp.*, vol. 29, pp. 396-406, 1975.
- [8] Carpenter, M. H., D. Gottlieb, S. Abarbanel, "The stability of numerical boundary treatments for compact higher-order finite difference schemes," *ICASE*, ICASE Rept. # 91-71, 1991.
- [9] Shang, J. S., "Characteristic-based algorithms for solving Maxwell's equations in the time-domain," *IEEE Ant. Propagat. Mag.*, vol. 37, pp. 15-25, 1995.
- [10] Cangellaris, A. C., "Numerical stability and numerical dispersion of a compact 2D-FDTD method used for the dispersion analysis of waveguides," *Microwave Theory and Guided Wave Letters*, vol. 3, pp. 3-5, 1993.
- [11] Gaitonde, D., J. L. Young, and J. S. Shang, "Towards the construction of fourth-order difference schemes for transient EM wave simulation: Collocated Grid Approach," (In preparation).
- [12] Mur, G., "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic-field equations," *IEEE Trans. Electromagn. Compat.*, vol. 23, pp. 377-382, 1981.
- [13] Gottlieb, D., and B. Yang, "Comparisons of staggered and non-staggered schemes for Maxwell's equations," *12 Annual Rev. of Progress in Appl. Comp. Electromagn.*, Monterrey, CA, pp. 1122-1131, 1996.
- [14] Thomas, L. H., "Elliptic problems in linear difference equations over a network," *Watson Sci. Comput. Lab. Rept.*, Columbia University, New York, 1949.
- [15] Press, W. H., B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, 1986.

- [16] Gaitonde, D. and J. S. Shang, "High-order finite-volume schemes in wave propagation phenomena," *27th AIAA Plasmadynamics and Lasers Conf.*, AIAA 96-2335, June 1996.
- [17] Young, J. L., "A higher-order FDTD method for EM propagation in a collisionless, cold plasma," *IEEE Trans. Ant. Propagat.*, September, 1996.
- [18] Fyfe, D. J., "Economical evaluation of Runge-Kutta formulae, *Math. Comput.*, vol. 20, pp. 392-398, 1966.
- [19] Williamson, J. H., "Low-storage Runge-Kutta schemes," *J. Comp. Physics.* vol. 35, pp. 48-56, 1980.
- [20] Wait, J.R., *Introduction to Antennas and Propagation*, Peter Peregrinus Ltd, London, UK, 1986.